



Evaluation and Application of Space Telescope
Aberration Sensing Using Phase Diversity

DISSERTATION

David J. Lee
Captain, USAF

AFIT/DS/ENP/97-05

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

DTIC QUALITY INSPECTED 1

19970430 007

AFIT/DS/ENP/97-05

Evaluation and Application of Space Telescope Aberration Sensing Using Phase
Diversity

DISSERTATION

David J. Lee
Captain, USAF

AFIT/DS/ENP/97-05

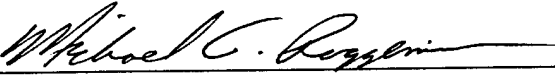
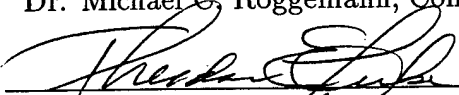
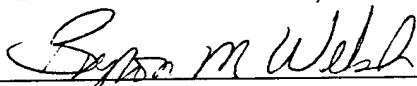
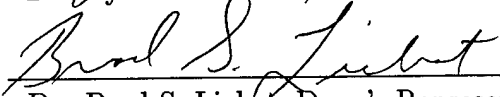
Approved for public release; distribution unlimited


EVALUATION AND APPLICATION OF SPACE TELESCOPE
ABERRATION SENSING USING PHASE DIVERSITY

David J. Lee, B.S., M.S.

Captain, USAF

Approved:

	<u>4-2-97</u>
Dr. Michael C. Roggemann, Committee Chairman	
	<u>4-2-97</u>
Dr. Theodore E. Luke, Committee Member	
	<u>4-2-97</u>
Dr. Byron M. Welsh, Committee Member	
	<u>4-2-97</u>
Dr. Brad S. Liebst, Dean's Representative	


Dr. Robert A. Calico, Jr., Dean,
Graduate School of Engineering

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

AFIT/DS/ENP/97-05

EVALUATION AND APPLICATION OF SPACE TELESCOPE
ABERRATION SENSING USING PHASE DIVERSITY

DISSERTATION

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

David J. Lee, B.S., M.S.
Captain, USAF

March, 1997

Approved for public release; distribution unlimited

Acknowledgements

First and foremost, I wish to thank my wife, for agreeing to stay for another 3 years in “sinus valley”, and for supporting our decision to pursue an AFIT doctorate at every step of the way. This support often took the form of her doing things that I should have been doing, but didn’t have the time to accomplish, including running a household while I was sitting in my cubicle at AFIT!

I’m sure that Maj. Roggemann was instrumental to the acceptance of a space operations master’s student into the physics department’s doctoral program. I appreciate the vote of confidence, as well as the continual support in terms of academic expertise, research advice, and the nuts and bolts aspects of how to be both a scientist and an Air Force officer. Similarly, I am grateful to the rest of my committee for their help in reviewing this document, as well as previous papers and research projects. Byron Welsh’s excellent instruction in such difficult topics as Statistical Optics proved invaluable, and was a key factor in my understanding of optics and imaging. I am also indebted to fellow students who helped with such things as the minutia of complicated mathematical derivations or the ins and outs of L^AT_EX or Unix.

I finally wish to congratulate AFIT for the existence of an excellent graduate program in the form of the Engineering Physics doctorate. The program is unique in that students like myself were allowed to straddle the diverse disciplines of Physics and Electrical Engineering. I feel that this interdisciplinary approach will prove invaluable to my future career as an Air Force researcher.

David J. Lee

Table of Contents

	Page
Acknowledgements	iv
List of Figures	ix
List of Tables	xi
Abstract	xii
 I. Introduction and overview	 1
1.1 Problem overview	1
1.2 Phase diversity in space telescopes	2
1.3 Chapter descriptions	5
1.4 Key results	7
 II. Feasibility of using WFS data in space telescope image deconvolution	 10
2.1 Introduction	10
2.2 Basic imaging relations	11
2.2.1 Linear systems imaging model	11
2.2.2 Aberrations and OTFs	12
2.2.3 Zernike polynomial decomposition	14
2.3 Monte-Carlo feasibility study	16
2.3.1 Deconvolution from wavefront sensing	18
2.3.2 Deconvolution estimators	20
2.4 Simulation	22
2.4.1 Phase screen generation	24
2.4.2 Randomness considerations	25

	Page
2.4.3 Simulated wavefront sensing	26
2.5 Simulation results	28
2.5.1 Signal-to-Noise Ratio	29
2.5.2 OTF phasor angle	30
2.5.3 Effective Strehl ratios	32
2.6 Conclusion	37
III. Phase diversity wavefront sensing: theoretical considerations	38
3.1 Least-squares phase diversity wavefront sensing	38
3.2 Motivation for collecting diversity images	44
3.2.1 Object independence	44
3.2.2 Uniqueness and ambiguity in image-based WFS	46
3.3 Stabilization of the inverse problem	51
3.3.1 Weighted least-squares noise suppression	51
3.3.2 Demonstration of regularization	55
3.4 Conclusion	57
IV. Monte-Carlo analysis of least-squares phase diversity: space telescope scenario	58
4.1 Introduction	58
4.2 Space telescope aberration bases	60
4.3 Computer Simulation	64
4.3.1 Random aberration generation	64
4.3.2 Simulation parameters	66
4.4 Simulation Experimental Results	69
4.4.1 Overall pupil-domain RMS WFS error	69
4.4.2 Zernike mode WFS errors	78
4.4.3 Using PDWFS estimates in phase deconvolution	79
4.5 Conclusion	93

	Page
V. Cramér-Rao analysis of phase diversity imaging	95
5.1 Introduction	95
5.2 Notation and preliminary expressions	96
5.2.1 Fourier optics revisited	96
5.2.2 Preliminary expressions: partial derivatives	98
5.3 Phase diversity Cramér-Rao lower bound derivation	100
5.3.1 Cramér-Rao concept	100
5.3.2 Phase diversity semi-classical photodetection model	101
5.3.3 Phase diversity Fisher information matrix	103
5.4 Simulation examples	107
5.4.1 Optimal defocus example	109
5.4.2 Curvature sensing example	112
5.4.3 Other experiments	115
5.5 Conclusion	121
VI. Applications of phase diversity to ground-based adaptive optical systems: Diagnosing image-path-only aberrations	124
6.1 Problem introduction	124
6.2 Adapting phase diversity to this problem	126
6.3 Example application	129
6.4 Remarks on practical application of the technique	137
6.5 Conclusion	140
VII. Conclusion	142
7.1 Summary of results	142
7.2 Follow-on research ideas	144
Appendix A. Remarks on pupil weighting and PSF normalization	146
Bibliography	147

	Page
Vita	152

List of Figures

Figure	Page
1. Simplified schematic diagram of one possible phase diversity setup. .	3
2. Simplified block diagram of the deconvolution from wavefront sensing technique.	19
3. Computer simulation simplified block diagram.	23
4. OTF Signal-to-Noise Ratios, PCWFS cases.	31
5. OTF phasor angle errors, PCWFS cases.	33
6. Strehl Ratios, PCWFS cases.	35
7. Conceptual diagram of the general phase retrieval uniqueness problem.	47
8. Phase ambiguity demonstration with spherical aberration.	49
9. Objective functions for spherical aberration ambiguity demonstration.	50
10. Example cone filter.	53
11. Demonstration of the effectiveness of noise-rejection.	56
12. Possible configuration of the proposed Next Generation Space Telescope.	62
13. Simplified approximation of the NGST.	63
14. Computer simulation simplified block diagram.	65
15. Simulated mean-squared aberration parameters.	68
16. Pupil-averaged RMS phase diversity WFS estimation errors, Zernike cases.	71
17. Pupil-averaged RMS phase diversity WFS estimation errors, NGST cases.	72
18. Ensemble-averaged, pupil-averaged RMS phase diversity WFS estimation errors, Zernike cases.	76
19. Ensemble-averaged, pupil-averaged RMS phase diversity WFS estimation errors, NGST cases.	77
20. Estimation accuracy on a Zernike mode-by-mode basis.	80

Figure	Page
21. OTF Signal-to-Noise Ratios, Zernike PCPD case 1.	83
22. OTF Signal-to-Noise Ratios, Zernike PCPD case 2.	84
23. OTF Signal-to-Noise Ratios, NGST PCPD case 1.	85
24. OTF Signal-to-Noise Ratios, NGST PCPD case 2.	86
25. OTF phasor angle errors, Zernike PCPD case 1.	88
26. OTF phasor angle errors, Zernike PCPD case 2.	89
27. OTF phasor angle errors, NGST PCPD case 1.	90
28. OTF phasor angle errors, NGST PCPD case 2.	91
29. Average point source image slices before and after PCPD.	92
30. Graphical explanation of the procedure used to obtain Cramér-Rao figure-of-merit metric values.	110
31. Optimal defocus CRLB data.	111
32. Schematic comparison of the curvature sensing and phase diversity techniques.	113
33. Optimal curvature sensing CRLB data.	114
34. CAD rendering of a Russian Ocean Reconnaissance satellite.	119
35. Simplified block diagram of an adaptive optics imaging system with image path aberration.	125
36. Schematic for using phase diversity to diagnose an AO image aberration that is not present in the WFS optics path.	128
37. Example images frames from ensemble of SOR image and WFS data.	129
38. Average star image frame and its corresponding original Hartmann WFS-based PSF estimate	131
39. Image-based and WFS-based average Fourier transforms.	133
40. The phase diversity estimate of the image-path aberration, restricted to Zernike modes 4-22. Gray-level map is shown.	136
41. Average phase-diversity-augmented, WFS-based PSF estimates, 19-mode case.	137
42. Image-based and WFS-based average Fourier transforms.	138

List of Tables

Table		Page
1.	The effects of PCWFS on Strehl Ratio.	36
2.	The summary of results of the various phase diversity CRLB experiments. Photocounts normalized to unity.	115
3.	Noll-modified Zernike coefficients of the diagnosed image-path aberration phase screen, 19-mode estimate.	135

Abstract

Due to mechanical aspects of fabrication, launch, and operational environment, space telescope optics can suffer from unforeseen aberrations, detracting from their intended diffraction-limited performance goals. This dissertation gives the results of simulation and theoretical studies designed to explore how wavefront aberration information for such “nearly diffraction-limited” telescopes can be estimated via the Gonsalves (least-squares) phase diversity technique.

Simulation studies utilized numerically simulated imaging models of both monolithic and segmented space telescope mirrors. The segmented case is a simplified model of the proposed Next Generation Space Telescope (NGST). The Monte-Carlo simulation results quantify the accuracy of phase diversity as a wavefront sensing (WFS) technique in estimating the pupil phase map. Simulation results give an indication of the minimum light level required for reliable estimation of a large number of aberration parameters under the least-squares paradigm. For weak aberrations averaging 0.10λ RMS, the average WFS estimation errors obtained here range from a worst case 0.057λ RMS to a best case of only 0.002λ RMS, depending upon the light level as well as the types and degrees-of-freedom of the aberration present. These studies are unique in their incorporation of photon statistical considerations.

Theoretical investigation of space telescope phase diversity imaging consisted of Cramér-Rao lower bound analysis. The CRLB expressions given here provide a novel computational tool for assessing the merits of particular phase diversity imaging configurations. One key result of such an analysis is the proposal that phase diversity WFS estimation might, under certain conditions, be carried out using symmetrically defocused images, as in the curvature sensing technique. In the

test cases demonstrated here, such a symmetrically defocused configuration resulted in smaller minimum mean-squared estimation errors.

Phase diversity was also applied to the estimation of fixed optical aberrations in an operational adaptive optics system. Nineteen Zernike modes of an aberration that was present in the image path of an operational adaptive optics system were successfully estimated.

EVALUATION AND APPLICATION OF SPACE TELESCOPE ABERRATION SENSING USING PHASE DIVERSITY

I. Introduction and overview

1.1 Problem overview

Optical aberrations degrade the imaging performance of optical telescopes. In response, adaptive optical and space-based telescopes have been developed which can greatly reduce the effects of aberration from sources external to the telescope, such as the turbulent atmosphere in the case of astronomical imaging. These efforts have brought optical astronomical imaging into the realm of nearly diffraction-limited performance. It is obvious, however, that adaptive optics systems cannot respond with infinite speed or perfect accuracy, and that space telescopes can suffer from unforeseen aberrations, such as the infamous spherical aberration introduced into the fabrication of the Hubble Space Telescope (HST) (7, 9, 18). Is there some technique we can use to diagnose these “residual”, weak aberrations?

One candidate technique, which is the focus of this dissertation, is known as the phase diversity aberration sensing technique—an *a posteriori* image-based wavefront estimation method which was first presented for optical imaging by Gonçalves (26). The term “*a posteriori*” is used here to denote the fact that phase diversity is not generally a real-time wavefront sensing technique, but instead depends on post-processing of recorded image data. The technique can be contrasted against traditional wavefront sensor (WFS) devices which measure pupil-related quantities more directly. One such type of traditional WFS device is the Hartmann slope sensor

that is often found in adaptive optics (AO) systems. In ground-based AO imaging systems aberration estimates need rapid updating in near-real-time—on the order of milliseconds—in order to allow for compensation of turbulence-induced aberrations (4, 32, 72, 77, 80). But such traditional WFS devices are in sharp contrast with the phase diversity technique. The phase diversity WFS method is more akin to the phase retrieval (17, 25) or blind deconvolution (5, 40, 46) families of post-processing techniques, a fact which will be discussed in subsequent chapters.

The phase diversity approach is motivated by the fact that the mathematical mapping from the set of all possible pupil phase screens to the set all possible point-spread functions (PSFs) is a many-to-one mapping. In order to invert this mapping, which is the goal of image-based aberration sensing, the phase diversity methodology requires simultaneous collection of multiple images, each formed via a slightly different pupil phase screen, such as in the simplified example configuration shown in figure 1. In that example, the second, defocused image is considered to be the “diversity” image, and the difference between the two pupil phase screens would be a quadratic, consistent with the physical defocus of the optical hardware setup. Phase diversity post-processing then consists of fitting the “best” unknown aberration estimate to this set of multiple, “phase-diverse” image measurements. This fitting is accomplished via the minimization of some appropriate cost function. Recent example applications of the phase diversity technique can be found in references (2, 8, 23, 49, 63).

1.2 *Phase diversity in space telescopes*

Even though free of atmospheric turbulence effects, space-based telescopes can still suffer from their own fixed or quasi-static internal aberrations. These aberrations can originate from various thermal/mechanical stresses on the optics, or errors in

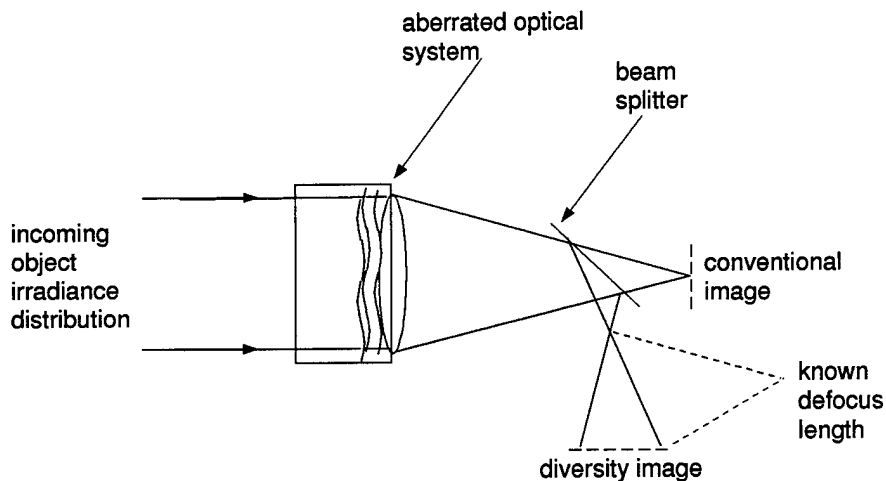


Figure 1. Simplified schematic diagram of one possible phase diversity setup.

the fabrication of the optical system (48). Rapidly changing aberrations may also arise due to vibrations induced by spacecraft components such as reaction wheels or solar array drive motors (57). If knowledge of these pupil aberrations is somehow made available by way of some sort of wavefront sensing (WFS) technique, then the aberrations can be at least partially compensated through either active optical components (38, 53, 73, 77) or post processing (20, 65, 81).

In this dissertation an analysis of the phase diversity method is combined with the space-telescope imaging problem. We connect these two topics because the phase diversity WFS technique offers a number of potential practical advantages over traditional WFS methods in the case of space-based imaging. For example, the hardware implementation of the phase diversity technique is relatively straightforward, in comparison to the complicated optical hardware systems of standard Hartmann sensors and shearing interferometers. These traditional WFS systems are also themselves subject to misalignment, misregistration, and aberration errors. Moreover, for telescopes with segmented primary mirrors, such as the proposed Next Generation Space Telescope (NGST), discussed in chapter 4 and refs. (1, 13, 79, 54, 14), conventional

slope sensors would be ill-suited to measuring the sharp pupil phase discontinuities introduced by misaligned segments (10).

The optical simplicity that phase diversity WFS offers is especially relevant to the space telescope problem, since spaceborne systems operate in a hostile, inaccessible environment, and the systems designers face significant cost, size, and weight constraints (3). Additionally, the phase diversity methodology is relatively immune to any systemic WFS system errors that may manifest themselves once that spacecraft is in orbit, since the technique uses the target object as a WFS reference (62). Moreover, under the least-squares formalism developed in chapter 3, the phase diversity technique works regardless of the object being viewed—a bright, point-like reference beacon is not specifically called for. Traditional Hartmann and shearing-interferometric WFS systems, on the other hand, generally do require views of unresolved point source beacons to operate. The price paid for these phase diversity advantages is that phase diversity wavefront sensing (PDWFS) measurements require extensive post-processing in order to extract and reconstruct the aberration information, as opposed to the simple matrix multiplications used for reconstruction of Hartmann wavefront sensor measurements (62, 72).

The primary goal of this dissertation is to show the development and implementation of methods for quantifying the performance of phase diversity wavefront sensing. Simulation analyses of phase diversity performance are given, constrained to various space-telescope imaging models of interest, namely, systems with either:

1. a monolithic mirror experiencing a quarter-wave RMS aberration or less, or
2. a similarly aberrated segmented mirror, approximating the proposed configuration of the Next Generation Space Telescope discussed above.

Theoretical analysis of phase diversity problem is given in terms of fundamental estimation accuracy limits. Finally, an example of a unique practical application of phase diversity is demonstrated for the case of aberration diagnosis in an operational adaptive optics telescope system.

1.3 Chapter descriptions

Before proceeding with an in depth analysis of some particular WFS methodology such as phase diversity, the following question should be addressed: “would WFS information actually be of any significant value for a nearly-diffraction-limited space telescope?” To that end, chapter 2 presents the results of a Monte-Carlo feasibility study of deconvolution from wavefront sensing (DWFS), simulating a generic, unspecified, Zernike mode WFS with user-definable estimation performance. Experiments are carried out using an imaging simulation, to determine if anything useful can be done with noisy, imperfect WFS measurements in terms of post-processing of weakly aberrated images. These numerical experiments show that, given such WFS information, imaging resolution performance can be improved noticeably via phase deconvolution, which is also defined and discussed in that chapter.

This positive result for a generic, unspecified WFS gives reason to believe that the specific PDWFS method, combined with deconvolution processing, may also prove similarly successful. Chapter 3 proceeds with a theoretical discussion of the phase diversity technique, as presented originally by Gonsalves (25, 26). This chapter also discusses the practical motivation for collecting multiple, diverse images, as opposed to a single image. The topic of inverse problem regularization is also covered.

Chapters 4 and 5 provide an in-depth investigation of the performance of the specific Gonsalves phase diversity technique itself, first in terms of numerical, Monte-Carlo space telescope imaging simulations, chapter 4, and then in terms of fundamental, estimation-theoretical performance limits, chapter 5, for various space telescope scenarios of interest.

The Monte-Carlo analysis of optical aberration and misalignment sensing via phase diversity (chapter 4) is a research task that was first suggested in reference (61). However, such an analysis of PDWFS has never been given in any subsequent published literature, making this chapter a unique contribution to the field. The analysis in this chapter is especially relevant due to the fact that one group of the pupil simulations closely models a proposal for the Next Generation Space Telescope (NGST), a proposed follow-on to the HST that is currently on-orbit.

The fundamental estimation-theoretic performance analysis of PDWFS, given in chapter 5, is likewise a generalization of previous work—specifically, the estimation-theoretical analysis of the Hubble Space Telescope problem developed in reference (18). But reference (18) tackles only the single-image phase retrieval problem, for point-source imaging; whereas the work shown in chapter 5 generalizes the estimation-theoretic analysis to include *phase diversity* imaging of *any* given target source. The work shown here appears to be the first to take the extra step of actually applying Cramér-Rao analysis to the the phase diversity problem, and the first to address some of the interesting questions that can be attacked using numerical analysis of the derived Cramér-Rao expressions.

Chapter 6 details how the phase diversity technique could also have relevant application in the realm of ground-based adaptive optical imaging systems. First a novel re-interpretation of the the phase diversity methodology is presented, show-

ing how it can be used to diagnose an image-path-only aberration in an operational adaptive optical system—an aberration that is not sensed by the Hartmann wavefront sensor. This idea is demonstrated using actual imagery and WFS data from an operational adaptive optics (AO) system, the USAF Phillips Laboratory Starfire Optical Range. The results shown there represent a potentially useful addition to the collection of tools available to the practicing astronomer who is using adaptive optical systems. The results are also significant in that they represent a validation of phase diversity methodology using real imagery.

Chapter 7 presents relevant conclusions and recommendations for further research. The overall conclusion is that a space telescope system would be well served by the integration of the phase diversity aberration sensing technique into its operational concept. The phase error estimates could be used to mechanically correct the phase errors, through adaptive optics if appropriate or possible. Alternatively, the phase error estimates could be integrated into a post-processing methodology, such as the Fourier phase filtering technique demonstrated in this dissertation. Ideas for follow-on research include the possibility of using a weighted least-squares methodology for including model information into the phase diversity technique. Other research projects include numerical simulation of other space telescope aberration models, besides the examples given here. Similarly, the Cramér-Rao numerical evaluations given in chapter 5 could be replicated for a large variety of phase diversity configurations and experiments not evaluated here.

1.4 Key results

The significant experimental and theoretical results of this dissertation are summarized below.

1. The phase filtering technique of image restoration was validated for the case of nearly diffraction-limited, noiseless, point-source imaging. Even when using very noisy WFS data, WFS SNR = 2, to estimate OTF phases, a quantifiable improvement in deconvolved imagery was noted. For example, in a simulated case of 0.10λ RMS aberration, the noise-effective cutoff frequency is increased from 0.8 to 0.9 normalized spatial frequency.
2. In-depth numerical experimentation on the PDWFS technique for the case of space telescope point-source imaging showed that accurate pupil estimates can be obtained even under very low light conditions. One experimental case, for example, involved PDWFS estimation of 0.10λ RMS piston errors for a segmented space telescope model from dim, photon-limited, point-source images ($\overline{K} = 1000$). These simulation experiments yielded phase estimates with average RMS errors of 0.012λ . This represents a breakthrough, since such piston misalignments of a segmented telescope could not be estimated using standard slope sensor systems.
3. A significant limitation to the example case discussed in item 2 above was noted. The experiment in question here dealt again with the simulated model of a segmented space telescope, with similar PDWFS estimates being made for *both* segment piston and tilt misalignment errors. Under the same low-light conditions as before, 20% of these higher degree-of-freedom Monte-Carlo PDWFS experimental cases ended in failure, the algorithm converging on WFS estimates that were incorrect by several orders of magnitude. One observation that may be consistent with these failure outcomes is the fact that Gonsalves PDWFS does not account for photon noise in a maximum-likelihood sense. Items 2 and 3 summarize the first such attempt at analysis of the photon-noise limitations of the Gonsalves technique.

4. A unique Cramér-Rao lower bound analysis, presented here for general, photon-limited, phase diversity imaging, provides a novel computational tool for assessing the merits of particular phase diversity imaging configurations. One key result of this analysis is the proposal that phase diversity WFS estimation might be more appropriately carried out using symmetrically defocused images, along the same lines as the images collected in the curvature sensing technique. A variety of other modifications to the standard PDWFS implementation are proposed and investigated via CRLB analysis. Furthermore, it is shown that the mean-squared error observed in previous PDWFS Monte-Carlo experiments approached their appropriate CRLBs to within a factor of 2.
5. Finally, an innovative re-interpretation of the phase diversity technique is shown to yield a powerful new technique for the estimation of fixed, image-path aberrations in an operational adaptive optics astronomical imaging system. The estimation of 19 Zernike modes of such an aberration is demonstrated using actual astronomical imagery provided by the USAF Phillips Laboratory Starfire Optical Range AO system.

II. Feasibility of using WFS data in space telescope image deconvolution

2.1 Introduction

Before proceeding with an in depth analysis of the phase diversity wavefront sensing (PDWFS) technique, this chapter addresses the question of whether or not wavefront sensor (WFS) information would actually be of significant value under a hypothetical, nearly-diffraction-limited space telescope imaging scenario. This chapter details a Monte-Carlo feasibility study of a generic, unspecified, Zernike mode WFS. The use of a “phase only” deconvolution estimator, first mentioned by Fried (20) is justified for the case of weakly-aberrated imaging. In order to find an upper bound on resolution performance, numerical simulation of noiseless, point-source imaging and deconvolution is performed. The goal of this chapter is to quantify the improvement gained by post-processing weakly aberrated images, given various levels of aberration and various levels of WFS accuracy. The overall conclusion is that imaging resolution performance can be improved noticeably via WFS-supported phase deconvolution, even for relatively inaccurate WFS measurements. This positive result—for a first order, proof-of-concept analysis with an unspecified WFS — shows that PDWFS, combined with deconvolution processing, may prove similarly useful in improving imaging resolution.

The remainder of this chapter is organized as follows. After reviewing basic Fourier optics, the deconvolution from wavefront sensing (DWFS) paradigm is explained. Then the simulation feasibility experiments for this chapter, and representative results from these experiments, are presented and discussed.

2.2 Basic imaging relations

This section reviews the linear systems imaging model and establishes the requisite mathematical notation. Fourier optics principles are used to derive the relation between wavefront phase aberrations and the image domain quantities.

2.2.1 Linear systems imaging model. Isoplanatic, or shift-invariant (LSI) incoherent imaging will be assumed throughout this dissertation. In the noiseless limit, isoplanatic image formation is modelled by a two-dimensional convolution operation (28),

$$(1) \quad i(\vec{x}) = o(\vec{x}) * h(\vec{x}),$$

where \vec{x} is a two dimensional position coordinate, $i(\vec{x})$ and $o(\vec{x})$ are image and object intensity distributions respectively, $h(\vec{x})$ is the system point spread function (PSF) or impulse response, and the asterisk, $*$, denotes the convolution operation,

$$(2) \quad \begin{aligned} c(\vec{x}) &= a(\vec{x}) * b(\vec{x}) \\ &= \int d\vec{x}_0 a(\vec{x}_0) b(\vec{x} - \vec{x}_0). \end{aligned}$$

Fourier transforming equation 1 gives

$$(3) \quad I(\vec{f}) = O(\vec{f})H(\vec{f}),$$

where \vec{f} is a two-dimensional spatial frequency coordinate, $I(\vec{f})$ and $O(\vec{f})$ are the image and object spectra respectively, and $H(\vec{f})$ is the optical transfer function

(OTF). The two-dimensional Fourier transform of some function $g(\vec{x})$ is defined as

$$(4) \quad G(\vec{f}) = \text{FT}[g(\vec{x})] = \int d\vec{x} g(\vec{x}) \exp[-j2\pi \vec{f} \cdot \vec{x}],$$

where the notation $\text{FT}[\cdot]$ is used to denote the Fourier transform operator. The inverse transformation from the spatial frequency domain back to the image domain is given as

$$(5) \quad g(\vec{x}) = \text{IFT}[G(\vec{f})] = \int d\vec{f} G(\vec{f}) \exp[j2\pi \vec{f} \cdot \vec{x}],$$

The notation $\text{IFT}[\cdot]$ is used to represent the inverse Fourier transform operator. Note that lower case symbols denote image-domain quantities, upper case symbols represent frequency domain quantities, and that $j = \sqrt{-1}$.

These quantities and relations are important because a linear, shift-invariant imaging system is completely specified by its impulse response, and therefore by its transfer function. The OTF is in turn dependent on the aberrations present in the imaging pupil, as reviewed next.

2.2.2 Aberrations and OTFs. Throughout this research optical aberrations are treated as thin, near-field phase screens. The thin phase screen assumption corresponds to the “thin lens” assumption of ray optics (29, 34), where a ray entering the screen at some transverse pupil coordinate exits the screen at essentially the same coordinate. The near-field assumption implies that the values of a pupil phase screen across the telescope aperture correspond only to the phase delays in the incoming optical wavefronts, with no amplitude scintillation effects (29).

In this effort it is generally presumed that undesired pupil phase delays are due to deformations and/or misalignment of the telescope mirror or associated optics,

which are in turn caused by vibration, mechanical flexing or bending of optical elements, or perhaps even optical fabrication and misalignment errors. The pupil phase aberration is denoted by $\phi(\vec{x})$, a quantity given in units of radians or waves (2π radians) of angular phase delay. The aberration $\phi(\vec{x})$ is the quantity which is to be estimated by a WFS procedure.

The basic Fourier optical principles found in refs. (28, 72) can be invoked to determine the effects of these aberrations on an LSI imaging system. If we denote the imaging aperture by the unity indicator function,

$$(6) \quad W(\vec{x}) = \begin{cases} 1 & \text{for } \vec{x} \in \text{pupil} \\ 0 & \text{otherwise,} \end{cases}$$

and define the generalized pupil function as

$$(7) \quad GPF(\vec{x}) = W(\vec{x})\exp[j\phi(\vec{x})],$$

then we can write the system OTF as

$$(8) \quad H(\vec{f}) = \frac{\text{ACF}[GPF(\vec{f}\lambda d_i)]}{\text{ACF}[GPF(\vec{f}\lambda d_i)]|_{\vec{f}=\vec{0}}}.$$

The symbol λ refers to the (center) wavelength of the (quasi-monochromatic) light being detected, d_i is the distance between the exit pupil and the imaging detector array, and the ACF $[\cdot]$ operator notation represents the autocorrelation function,

$$(9) \quad \text{ACF}[f(\vec{x})] = \int d\vec{x}' f(\vec{x}' - \vec{x}) f^*(\vec{x}').$$

The asterisk superscript denotes complex conjugation.

2.2.3 *Zernike polynomial decomposition.* It is often convenient to decompose a phase screen $\phi(\vec{x})$ into a linear combination of basis functions. When $\phi(\vec{x})$ is defined over a circular pupil of radius R , Zernike polynomials are often used:

$$(10) \quad \phi(Rr, \theta) = \sum_{i=1}^{\infty} a_i Z_i(r, \theta),$$

where r is the radial component on the unit disk (i.e. $|r| \leq 1$), and θ is the angular coordinate. The Zernike polynomials (6), $Z_i(r, \theta)$, as normalized and ordered by Noll (58), are given by:

$$(11) \quad \left. \begin{aligned} Z_{i=\text{even}}(r, \theta) &= \sqrt{n+1} R_n^m(r) \cos(m\theta) \\ Z_{i=\text{odd}}(r, \theta) &= \sqrt{n+1} R_n^m(r) \sin(m\theta) \end{aligned} \right\} m \neq 0,$$

and

$$(12) \quad Z_i(r, \theta) = R_n^0(r) \quad m = 0.$$

The radial polynomials $R_n^m(r)$ are in turn defined by

$$(13) \quad R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! [(n+m)/2 - s]! [(n-m)/2 - s]!} r^{n-2s}.$$

The indices m and n are non-negative integers such that $m \leq n$ and $(n-m)$ is even.

The Zernike polynomials are orthonormal, or,

$$(14) \quad \frac{\int d\vec{x} W(\vec{x}) Z_i(\vec{x}) Z_j(\vec{x})}{\int d\vec{x} W(\vec{x})} = \delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases}$$

where $\delta_{i,j}$ represents the Kronecker delta function. The coefficients, a_i , are therefore given by the inner product of $\phi(\vec{x})$ with the various Z_i 's (58):

$$(15) \quad a_i = \frac{\int d\vec{x} W(\vec{x}) \phi(\vec{x}) Z_i(\vec{x})}{\int d\vec{x} W(\vec{x})}.$$

Orthonormality also leads to the following useful property:

$$(16) \quad \frac{\int d\vec{x} W(\vec{x}) (\phi(\vec{x}))^2}{\int d\vec{x} W(\vec{x})} = \sum_{i=1}^{\infty} a_i^2,$$

where again, $\phi(\vec{x})$ is given in equation 10. Equation 16 is a generalized Parseval relation which states that the pupil-averaged root-mean-square (RMS) value of a phase screen is equal to the root-sum-squared value of the corresponding Zernike coefficients. This property is useful because aberrations and telescope "wavefront error budget" specifications are often given in terms of pupil-averaged RMS values. Zernike polynomials are widely used in the optics literature due to the correspondence of the first dozen modes with classical optical aberrations. For instance, Z_1 represents piston, Z_2 and Z_3 correspond to x and y tilt, and Z_{11} corresponds to spherical aberration. The first 3 Zernike modes will be ignored throughout this dissertation whenever the Zernike basis set is used to decompose an aberrating phase screen. The first mode Z_1 is neglected because imaging and WFS systems are insensitive to piston. Instantaneous image quality is unaffected by tilt, Z_2 and Z_3 , because they correspond to simple image displacement, but cause no other defect in the image.

Although the Zernike polynomials given above are defined over a unit circle, the aperture functions of many common reflector telescopes configurations are annular in nature, due to the central obscuration caused by a secondary mirror. It should

therefore be noted that corresponding Zernike polynomials, defined and orthonormalized over an annulus, also exist (18, 51). However, when Zernike polynomials are used in this dissertation, we will limit ourselves to consideration of filled circular pupils.

2.3 Monte-Carlo feasibility study

The well known problem of incoherent astronomical imaging through the Earth's atmosphere has motivated the development of a host of techniques intended to mitigate the effects of random optical aberrations. Most of these techniques can be grouped into pre-detection techniques and post-detection techniques. In pre-detection – or adaptive optics (AO) – compensation, pupil phase aberrations are sensed with a wavefront sensor (WFS). This WFS information is used to drive active optical components to *mechanically* cancel the wavefront phase aberrations *before* the image is detected (32). Post-detection compensation entails the recording of images for *computational* processing *after* the image is detected (74, 81). The goal of the post-processing is to exploit the mathematical and/or statistical properties of the imaging situation in order to computationally compensate for the effects of the aberrations (45). AO systems are expensive and hardware intensive, while post-processing techniques usually incur intensive computational requirements. The remainder of this chapter will be devoted to exploration of the use of a hybrid method, which uses principles from both of these traditional categories, and apply the method to a space-based imaging scenario.

Deconvolution from wavefront sensing (DWFS) refers to a relatively new class of aberration compensation techniques (81). As the name implies, this method involves the use of WFS information to sense the telescope pupil aberrations. But instead of using this information in real-time to drive AO hardware, the WFS data

are stored along with the corresponding image frames. Later, after the observing session, the aberration information is used in a post-processing scheme, such as deconvolution. DWFS may result in improved image quality, but without the substantial hardware requirements of pure AO systems or the heavy computational requirements of pure post-processing methods.

Post-detection, pre-detection, and DWFS techniques have all been analyzed for, and applied to the case of imaging in the presence of aberrations caused by atmospheric turbulence (55, 77, 72, 81). Similarly, post-detection (16, 18, 31, 82) and pre-detection (39, 53, 57, 83) techniques have been proposed and analyzed for the case of space-based telescope imaging. In this chapter we continue this parallel trend between ground-based and space-based imaging by extending the analysis of the DWFS technique mentioned above into a general simulated space-telescope scenario, with an eye towards the overall goal of evaluating a specific phase diversity wavefront sensing technique in the chapters which follow.

The primary difference – and advantage – of space-based imaging is the lack of aberrations due to atmospheric turbulence. But, even though free of these atmospheric turbulence effects, space telescopes can still suffer from their own aberrations (3, 7). Barring some catastrophic failure or miscalculation on the part of the telescope designers, these aberrations are usually expected to be orders of magnitude weaker than aberrations imposed by the atmosphere (16, 18, 57). However, even small aberrations can be significant if diffraction-limited performance was desired. With this in mind, the operative question for this section can be posed as: “What performance boost, if any, is gained in applying DWFS-style processing to space telescope imagery, given that the imagery is already ‘nearly’ diffraction-limited?”

If DWFS does indeed seem to promise measurable improvement in space telescope image quality, then several scenarios for exploiting the technique are possible. Space telescopes might, for example, suffer from unexpected aberrations due to fabrication errors, or mechanical aspects of the space environment or spacecraft operation. WFS-based post-processing might allow effective compensation of these aberrations. Image quality that would otherwise have been lost can be at least partially recovered, but without the cost of AO hardware or manned repair missions. Alternatively, the use of on-orbit aberration sensing in post-processing could permit designers to use less expensive optical manufacturing tolerances in the building of the telescope. Perhaps the image quality lost due to a less rigorous wavefront error budget could be made up through the use of DWFS.

The above considerations motivate this effort to quantify the performance of DWFS methods under non-atmospheric, weakly aberrated conditions.

2.3.1 Deconvolution from wavefront sensing. The purpose of a wavefront sensor (WFS) is to estimate the aberrating phase, $\phi(\vec{x})$, present in the effective entrance pupil of an incoherent imaging system. Typical WFS techniques involve viewing a reference beacon with wavefront slope sensors such as Hartmann or shearing interferometry devices in order to estimate the gradient of $\phi(\vec{x})$. An estimate of the phase screen, $\tilde{\phi}(\vec{x})$ can then be reconstructed by fitting basis functions to the integrated slopes by least-squares fit (80, 72). More recently, as mentioned in chapter 1, phase retrieval methods such as least-squares phase diversity estimation are being investigated for use in aberration sensing (8, 25). This particular method, for example, has the advantage that a spatially restricted beacon and complex optical hardware systems are not required, two significant advantages for space telescope systems, an assertion that will be discussed in chapter 3.

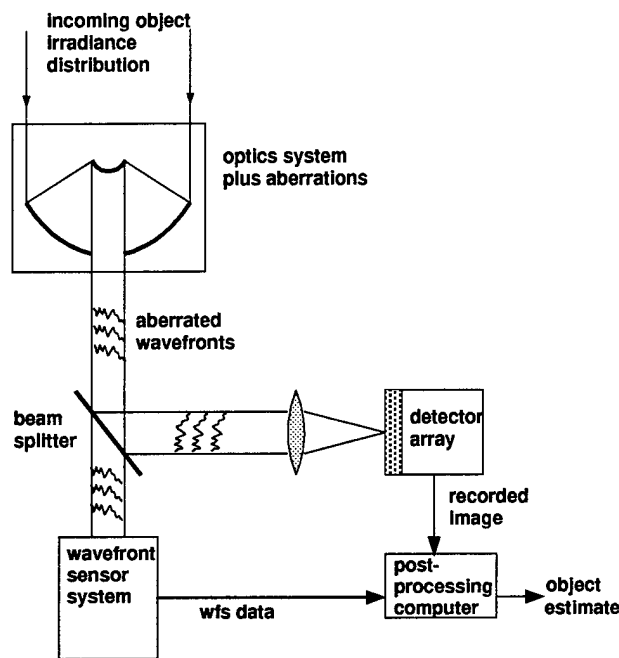


Figure 2. Simplified block diagram of the deconvolution from wavefront sensing technique.

In the DWFS image reconstruction technique, this WFS information is recorded for later processing, along with corresponding image frames, as shown in figure 2. With an eye towards the phase diversity technique, discussed in the chapters which follow, note that the “wavefront sensor system” block of this block diagram is not confined only to “traditional” WFS devices such as a Hartmann slope sensor array. The WFS aspect of the diagram could stand for any auxiliary information that has been gathered for WFS purposes. In the case of phase diversity, this auxiliary information would take the form of defocused, or diverse images that can be collected simultaneously via the beamsplitter arrangement, such as shown in figure 1 in chapter 1, and discussed in chapter 3. But since this chapter is serving as an initial feasibility study of the the DWFS technique in general, we will not restrict our consideration to any particular set of WFS characteristics.

2.3.2 *Deconvolution estimators.* Once the phase aberration estimate, $\tilde{\phi}(\vec{x})$, is determined through some WFS technique, the $\tilde{\phi}(\vec{x})$ quantity can then be substituted for $\phi(\vec{x})$ in equations 7 and 8. These relations yield an OTF estimate, $\tilde{H}(\vec{f})$. In their paper on the DWFS technique, Primot *et. al.* (65) use $\tilde{H}(\vec{f})$ in the following DWFS estimator:

$$(17) \quad \bar{O}(\vec{f}) = \frac{\langle I(\vec{f}) \tilde{H}^*(\vec{f}) \rangle}{\langle |\tilde{H}(\vec{f})|^2 \rangle},$$

where the tilde, \sim denotes an estimated quantity and the angle brackets indicate ensemble averaging of a set of realizations of an ergodic random process. Primot's estimator is a least-squares estimator under an LSI imaging model, as will be shown in chapter 3. If only a single realization is available, equation 17 reduces to

$$(18) \quad \begin{aligned} \bar{O}(\vec{f}) &= \frac{I(\vec{f}) \tilde{H}^*(\vec{f})}{|\tilde{H}(\vec{f})|^2} \\ &= \frac{I(\vec{f})}{\tilde{H}(\vec{f})}, \end{aligned}$$

commonly known as the inverse filter. Notice how the inverse filtering operation can also be written (suppressing \vec{f} dependence) as

$$(19) \quad \begin{aligned} \bar{O} &= \frac{I}{\tilde{H}} \\ &= O \frac{|H|}{|\tilde{H}|} e^{j(\Phi_H - \Phi_{\tilde{H}})}. \end{aligned}$$

where Φ_H and $\Phi_{\tilde{H}}$ denote the complex phasor angles of the H and \tilde{H} respectively. We see that this estimator involves a division of the OTF amplitudes, and a subtraction of their phasor angles.

Consider the quotient $|H|/|\tilde{H}|$. The denominator is generally an imperfect *estimate* of the numerator quantity. In the case under consideration here, for example, \tilde{H} is derived from noisy, reconstructed WFS data. For the spatial frequencies where this estimate has a low signal-to-noise-ratio, inverse filtering can result in amplification of noise, introducing unwanted variance into the deconvolved image estimate. In the worst case, if the estimated MTF modulus $|\tilde{H}|$ can take infinitesimally small values, there could be meaningless “infinite” values introduced into the DWFS estimate.

Various *ad hoc* “remedies” for this variance problem are discussed in (71) and in chapter 3. For images that already have sufficient high spatial frequency information content, it may be possible to sidestep the amplitude division issue by correcting only the OTF phase, leaving the OTF amplitudes unchanged. This idea was originally proposed by Fried (20):

$$\begin{aligned}
 \tilde{O}_{Fried} &= \frac{I\tilde{H}^*}{|\tilde{H}|} \\
 (20) \qquad &= O|H|e^{j(\Phi_H - \Phi_{\tilde{H}})}.
 \end{aligned}$$

In Fried’s formulation, the WFS-based OTF estimate is used to create a unit-phaser quantity which, in the limit of perfect wavefront-sensing, will provide the exact conjugate to the complex phase angle of the system OTF. The OTF modulus is not modified in this technique. Therefore, it has been suggested that this method would be useful for imaging systems where the OTF modulus already has sufficiently large value at high spatial frequencies, such as adaptive optics systems (68, 70) or space-based telescopes. For these reasons we adopt the Fried procedure of *phase-only correction from wavefront sensing* (PCWFS) in this dissertation. The performance

of this estimator on noiseless, point-source image measurements, given noisy WFS information, is analyzed by simulation in the next section.

We have shown that reducing the variance of an estimate is a *motivation* for phase-only filtering, but is there actual *justification* for neglecting modulus correction? References (33, 36, 59) all provide some insight into the relative importance of frequency domain phase information versus modulus information. Reference (33), for example, provides a mathematical theorem stating that a finite-extent discrete sequence is uniquely specified – to within a scale factor – by the phase of its Fourier transform, provided certain conditions on the z-transform of the sequence are met. Reference (59) provides informal discussion and heuristic examples supporting the idea that the frequency domain phase contains considerably more information about the visual content of an image than does the Fourier modulus. Finally, reference (36) presents results of statistical matched-filter experiments. Matched-filters which correlated image Fourier phase information of alpha-numeric characters were vastly more successful in providing detection in the presence of noise than were the corresponding Fourier-amplitude matched filters. These citations lend some theoretical weight to the phase deconvolution concept.

2.4 Simulation

A simplified flowchart of the imaging and reconstruction simulation is shown in figure 3. In this section the relevant details of the various flowchart elements are discussed.

The computer simulated pupil phases used here, and throughout this dissertation, are written as the arguments of complex exponentials to an $N \times N$ pixel array with a circular pupil aperture (equation 6), $N/2$ pixels in diameter, located at the

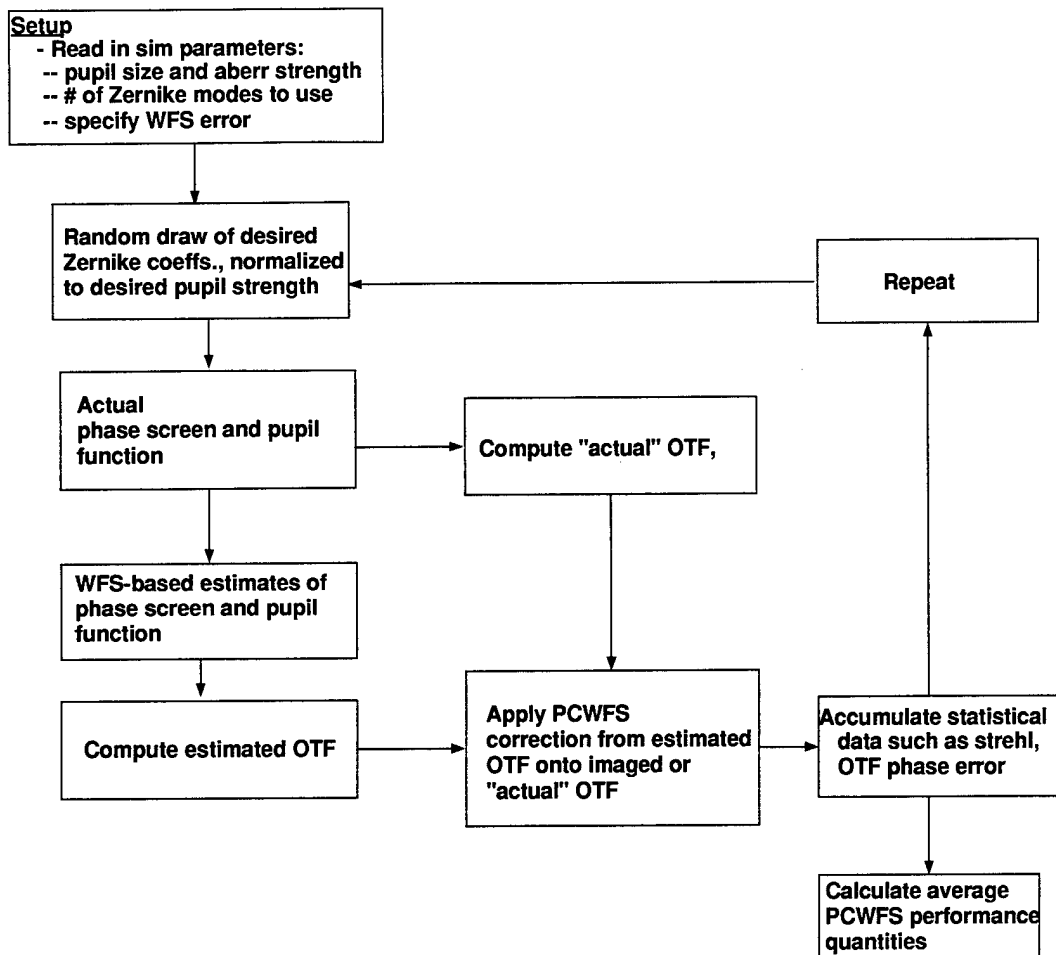


Figure 3. Computer simulation simplified block diagram.

center of the array. Using Fourier optical relations (section 2.2.1) and Fast Fourier transforms (FFTs), these phase screens will yield “Nyquist sampled” point-source images, with 2 pixels across a resolution cell, λ/D , with λ = center imaging wavelength, D = physical aperture diameter (72). The pupil phases for the simulations used in this section are specified in terms of the Zernike polynomials as ordered and normalized by Noll (58). In this section, all simulated point source images are noise free. Chapter 4 will show simulations which include photon noise, introduced

by using a random-number generation routine and the Poisson photon probability density described in chapter 5.

2.4.1 Phase screen generation. These simulation experiments incorporate the assumption that the RMS wavefront error budget of a general space telescope can take on values anywhere from on the order of 2% of a wave (39, 41, 52) up to approximately 30% of a wave (37), and that the aberrations are usually expected to be confined to the first two dozen Zernike modes, (16, 50, 57). An order-of-magnitude confirmation on the reasonableness of the assumptions above is provided by the well-known Hubble Space Telescope aberration. In that space telescope example, an aberration was erroneously introduced during fabrication. The aberration consisted of 50% of a wave RMS, mostly concentrated in Zernike mode 11 (7, 18, 56).

In this simulation scheme optical aberrations are modelled so that Zernike coefficients 4-22 are present. Piston and tilts are ignored as having no impact on the actual image quality of a single image realization. The remaining 19 coefficients are generated as independent, identically distributed, zero-mean Gaussian random variables. The pupil coefficients are all scaled by the same amount to give the desired ensemble-averaged, pupil-averaged RMS wavefront aberration.

From the 19 randomly drawn and appropriately scaled coefficients, a_i , a pupil phase screen is generated over a circular support:

$$(21) \quad \phi(Rr, \theta) = \sum_{i=4}^{22} a_i Z_i(r, \theta).$$

The simulated OTF is then calculated from this phase screen via equations 7 and 8.

2.4.2 *Randomness considerations.* The pupil Zernike coefficients generated by our simulation are statistically uncorrelated, so that we have

$$\begin{aligned} \langle a_i a_j \rangle &= \frac{\epsilon_{pupil}}{19} \delta_{i,j} \\ \text{for } i, j &\in \{[4, 5, \dots, 22]\} \end{aligned} \quad (22)$$

where $\delta_{i,j}$ is the Kronecker delta function, defined as

$$\delta_{i,j} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \quad (23)$$

Any non-zero correlation values between different Zernike modes (*i.e.* $i \neq j$) would imply *additional* statistical information about the structure of the aberrations. For the general space telescope problem such extra information is simply not available; often all that has been specified in the early proposal and design stages of some particular space telescope is the wavefront error budget.

No assumptions regarding the temporal nature or ergodicity of the phase screen random process are incorporated into this simulation model. Each realization of the phase screen might represent a temporally “fixed” aberration. Alternatively, each simulated phase screen may represent one sample of a temporally varying aberration, perhaps induced by mechanical vibrations of the spacecraft bus due to thermal flexing or spacecraft attitude corrections. In either case, the essential point is that the state of the aberration is unknown, and therefore random. The results given in section 2.5 represent ensemble averages taken across these random realizations of telescope pupils, as is appropriate for the type of feasibility study undertaken here.

In order for this simulation to be as general and as useful as possible, the desired RMS wavefront error budget is distributed randomly across Zernike modes 4 through 22. In this chapter, aberrations are modelled as being, on the average, equally distributed across all 19 of these Zernike modes, in a statistical sense. Of course, the actual aberration modal statistics of some given space telescope will be totally dependent on specifics such as how the mirror is built and mechanically supported. For example, at least one mechanical simulation study of a particular segmented space telescope mirror (57) revealed an approximate exponential falloff in aberration strength across the first 12 Zernike modes when the aberration is induced by mechanical vibrations of the satellite bus. A similar “modal falloff” model for Zernike aberration mode strength is adopted in chapter 4.

In order to provide a performance upper bound, these particular simulation experiments are limited to noiseless, point-source imaging, with no photon noise or read noise introduced into the images. Point source images are simulated since the performance of an isoplanatic imaging system is characterized by its PSF behavior.

2.4.3 Simulated wavefront sensing. Wavefront sensing is simulated by taking the known imaging pupil Zernike coefficients discussed in the previous subsection, that is the a_i values, and adding randomly drawn n_i noise coefficients to them. The error coefficients are also scaled so that ϵ_{wfs} , the ensemble-average of the pupil averaged WFS error, equals the desired value for the particular simulation run. From these corrupted Zernike coefficients an “estimated” phase screen and OTF are calculated:

$$(24) \quad \tilde{\phi}(Rr, \theta) = \sum_{i=4}^{22} (a_i + n_i) Z_i(r, \theta),$$

using the same symbol definitions as in equation 10 and 21. These random Gaussian WFS error coefficients n_i are used to account for such effects as WFS measurement noise, undersampling, and reconstruction error. The OTF, as measured by the simulated WFS, is then calculated from this phase screen via equations 7 and 8.

Notice that for this simulation, WFS estimation errors are, in an RMS sense, distributed equally across all of the 19 Zernike modes, regardless of order. In other words, the WFS system is simulated such that it will estimate Zernike coefficient a_4 or a_5 just as accurately, on the average, as it does a_{17} or a_{22} . This simulated approximation seems reasonable in light of results in reference (18), where an error analysis of space telescope aberration sensing is presented. Their results show that the lower bounds on Zernike coefficient estimation errors were of roughly the same order of magnitude for the coefficients out to Zernike mode 22 for a particular simplified space telescope model. The results of chapter 4 also suggest that, for the specific case of the phase diversity WFS technique, all of the first two dozen Zernike aberration modes are also simulated with about equal fidelity or accuracy, with no modes particularly favored over any other.

For convenience in describing the results which follow, let us define a "WFS SNR" quantity as

$$(25) \quad \text{WFS SNR} = \frac{\epsilon_{pupil}}{\epsilon_{wfs}}.$$

That is, the RMS wavefront sensor noise and error is expressed as a fraction of the RMS wavefront error budget. Recalling equation 24, the WFS SNR could be given

as

$$(26) \quad \text{WFS SNR} = \frac{\sqrt{\sum_{i=4}^{22} (a_i)^2}}{\sqrt{\sum_{i=4}^{22} (n_i)^2}}.$$

A WFS SNR of 2 means that the average RMS error of the WFS estimates is equal to one half the average RMS error of the actual aberration being estimated. Note that WFS SNR is a simple measure of simulated WFS fidelity, and is not to be confused with another definition of an “SNR” quantity given below for OTFs.

After the simulation has generated a truth-model OTF realization and the corresponding WFS-based estimate of that OTF, deconvolution performance statistics can be gathered. The results given in the next section represent averages of various performance metrics for implementation of the Fried PCWFS estimator.

2.5 Simulation results

The performance of the phase-only correction from wavefront sensing, or PCWFS, technique is given for a simulated space-based imaging scenario. Performance is specified in terms of OTF signal-to-noise-ratio, OTF phasor angle, and image domain Strehl ratios. These quantities are parameterized by various RMS wavefront error budget values and WFS error levels. Spectral quantities are given as radial averages, in terms of normalized spatial frequency. Radial averaging refers to a convenient method of condensing surface plots depending on 2-dimensional independent variables such as $\vec{f} = \{f_x, f_y\}$ down to line plots depending on the overall spatial frequency amplitude, $\rho = \sqrt{f_x^2 + f_y^2}$, by averaging the two-dimensional surface values around circles of constant radius ρ , centered on the origin. Spatial frequency

plotting variables are normalized to

$$(27) \quad \rho_{norm} = \frac{\rho\lambda}{D},$$

where D is the physical diameter of the telescope aperture, meaning that $\rho_{norm} = 1$ at the diffraction-limited cutoff.

2.5.1 Signal-to-Noise Ratio. The signal-to-noise ratio (SNR) of any random quantity is defined (29) as the expected value of the quantity divided by the standard deviation. The variance of a complex random process $X(\vec{f})$ is given by

$$(28) \quad \text{var} \{X(\vec{f})\} = \langle |X(\vec{f})|^2 \rangle - |\langle X(\vec{f}) \rangle|^2$$

where angle brackets denote the expectation value of random process, and so the SNR of this process is expressed as

$$(29) \quad \text{SNR}_X(\vec{f}) = \frac{|\langle X(\vec{f}) \rangle|}{\sqrt{\text{var} \{X(\vec{f})\}}},$$

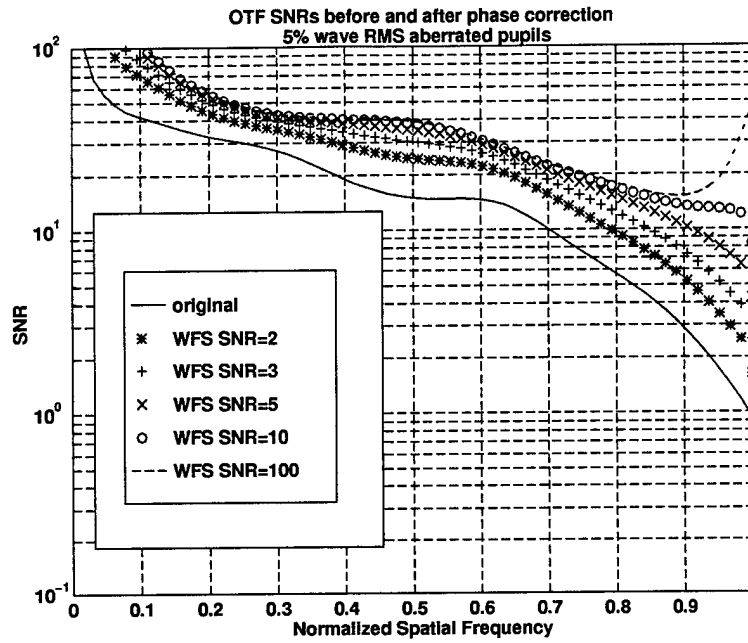
a real-valued function, which, for this analysis, will be applied to Fourier transform spectra, and thus will be a function of spatial frequency \vec{f} . Any unknown and/or randomly varying quantity or process introduced into a system model results in a decrease of the output SNR. The simulation model used here includes two sources of randomness in the imaging system:

- uncertainty about pupil aberrations (except for their RMS strength); and
- random wavefront sensor error.

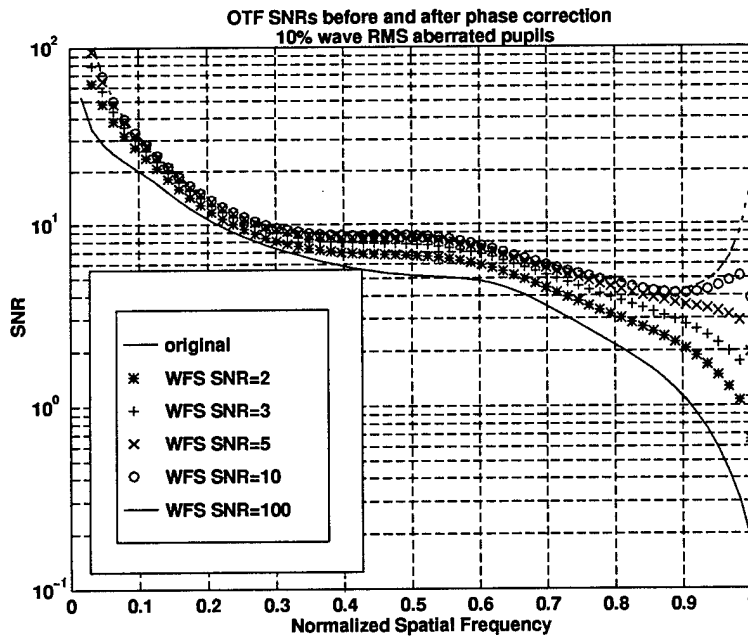
The relative impact of randomness on spatial frequency domain quantities is indicated by the solid-line curves shown in figure 4. These curves depict radially-averaged profiles of estimates of the quantity $SNR_H(\vec{f})$, the SNR of the effective optical transfer function before any sort of image reconstruction processing. The rest of the curves show the effect of PCWFS processing on the SNR of the effective system OTFs. The SNR quantities are simulation sample-based estimates, each obtained by averaging over 200 independent simulated pupil and WFS realizations for a variety of values of ϵ_{pupil} and ϵ_{wfs} .

An improvement in the OTF SNR is indicated even when very low quality WFS information is used. This improvement is evident in the two representative cases shown in figure 4, for pupils with RMS aberration strengths of (a) 0.05λ , and (b) 0.10λ . Consider, for example, the representative plot (a) of figure 4. That plot shows a PCPD improvement of at least a factor of 1.5 across all spatial frequencies out to cutoff, even when the WFS errors are half as strong as the aberrations being measured (WFS SNR = 2). This shows that rearranging image phases, even on the basis of relatively noisy WFS information, is still significantly better in terms of variance reduction than doing nothing at all. Similarly, the spatial frequency at which the average SNR drops below a threshold of 10 is shown to have been extended from 70% of cutoff out to 80% of cutoff, implying a corresponding improvement in average resolution (72).

2.5.2 OTF phasor angle. Recall that the goal of PCWFS is to drive the phasor angle of the effective imaging system OTF to zero in the limit of perfect OTF estimation by the WFS. Ideally, the effective OTF should be an entirely real-valued quantity, with the complex phasor angle of each spatial frequency component equal to zero. Figure 5 gives a detailed picture of the RMS phasor angle behavior



(a)



(b)

Figure 4. Radially-averaged profile of OTF Signal-to-Noise Ratios for some representative conditions both before and after PCWFS processing. The original, ensemble averaged, pupil averaged RMS aberration strengths were (a) 0.05λ , and (b) 0.10λ .

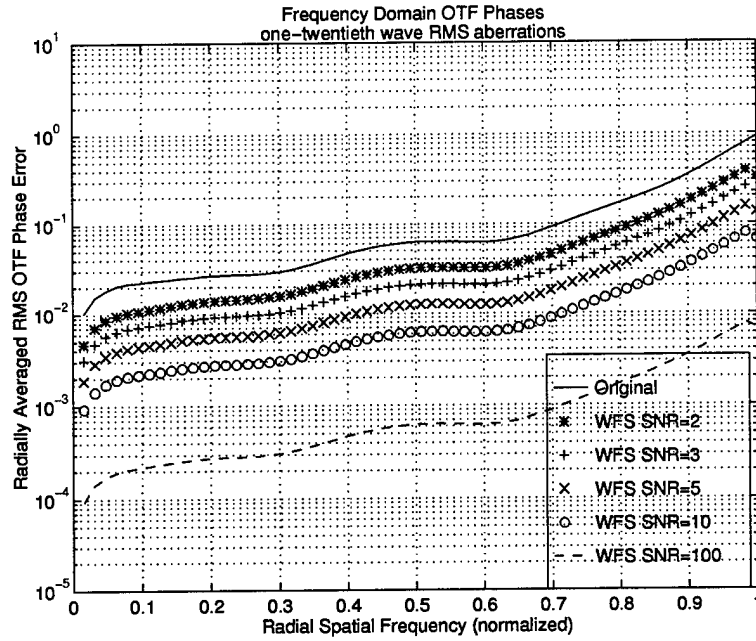
for the effective PCWFS OTF is at various spatial frequencies. Once again, these representative curves are given for wavefront errors of 5% and 10% of a wave RMS. Specifically, each curve in figure 5 represents

$$(30) \quad \text{curve}(\rho) = \text{Radial-average} \left\{ \frac{1}{200} \sqrt{\sum_{i=1}^{200} [\Phi_{H_i}(\vec{f})]^2} \right\},$$

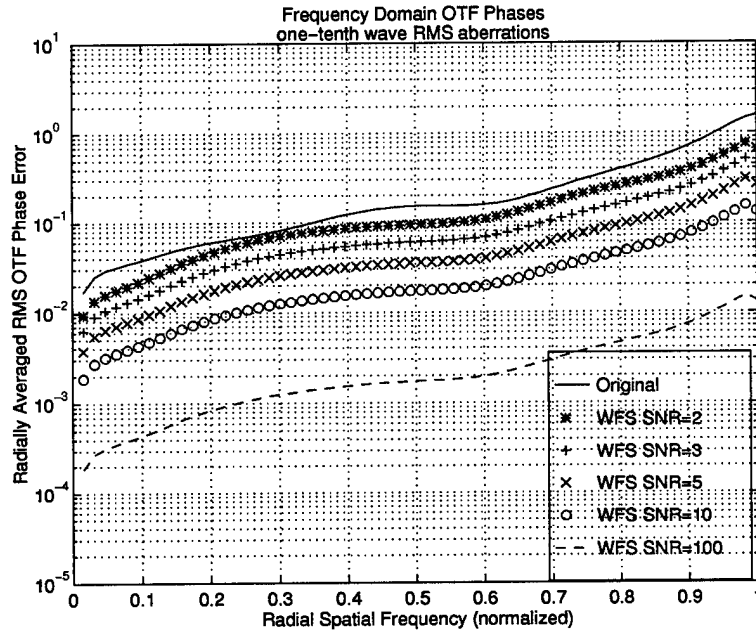
where $\rho = \sqrt{f_x^2 + f_y^2}$, the amplitude of a particular spatial frequency, and $\Phi_{H_i}(\vec{f})$ represents the phase angle of the particular OTF in question, at spatial frequency \vec{f} . In words, each curve is basically a radially-averaged profile of the RMS OTF phasor angle, as determined by averaging across 200 simulated realizations. This type of Fourier phase estimation error metric is commonly seen throughout the image reconstruction literature, *e.g.* ref. (69).

These representative phasor angle error plots show quantitatively that, as expected, the values of the effective PCWFS OTFs are driven towards the real axis. But more importantly, they also quantify the frequency domain impact of WFS estimation error on the PCWFS technique. As in the previous subsection, we again see that even the use of relatively noisy WFS data yields a significant improvement in imaging performance. When the WFS SNR is 2, the OTF phasor angles are generally reduced in amplitude by a factor of one-half. As expected, the lower curve of plot (a) shows that, given nearly perfect WFS information, we can correct the effective OTF such that it is essentially entirely real across all spatial frequencies.

2.5.3 Effective Strehl ratios. The spatial frequency domain performance metrics mentioned above are important considerations in terms of image reconstruction. But it may prove useful to tie these frequency domain results to more concrete effects in the image domain. The effects of PCWFS on the image domain are per-



(a)



(b)

Figure 5. Radial-average of RMS OTF phasor angles (radians). Original OTF values are compared to PCWFS effective OTF values with WFS SNRs of 2, 3, 5, 10 and 100, as shown. Each of the 200 simulated pupil realizations represented by each curve had (a) 0.05λ and (b) 0.10λ RMS pupil-averaged aberration.

haps most conveniently expressed in terms of effective Strehl ratio, which is defined as the ratio of the peak of the system PSF to the peak of the PSF of an unaberrated system. (28, 72).

Figure 6 shows a representative plot of the average effect the PCWFS processing has on the Strehl ratio of pupils with 0.10λ or less ensemble-averaged, pupil-averaged, RMS aberration. Each data point on the plot represents the average of 200 Strehl Ratios derived from 200 simulated realizations. Again, even for nearly diffraction-limited images, correcting the phase of the OTF using WFS information results in a quantifiable improvement in this traditional imaging performance metric. The WFS SNR was 10 in the PCWFS processed cases shown in figure 6.

The effective Strehl ratio data are also listed in table 1. Further insight can be gained on the effects of PCWFS on imaging by way of Marechal's approximation, where the Strehl Ratio, SR , is related to the pupil-averaged mean-square wavefront error, which is denoted by the symbol ϵ^2 , by the approximate relation

$$(31) \quad SR = \exp\{-\epsilon^2\},$$

a relation which holds for RMS wavefront errors up to about one-tenth of a wave (72). Notice that if P is defined as $P \equiv$ (percent-of-a-wave pupil-averaged RMS error), then

$$(32) \quad SR = \exp\left\{-\left(\frac{2\pi P}{100}\right)^2\right\}$$

and

$$(33) \quad P = \frac{50}{\pi} \sqrt{-\ln(SR)}.$$

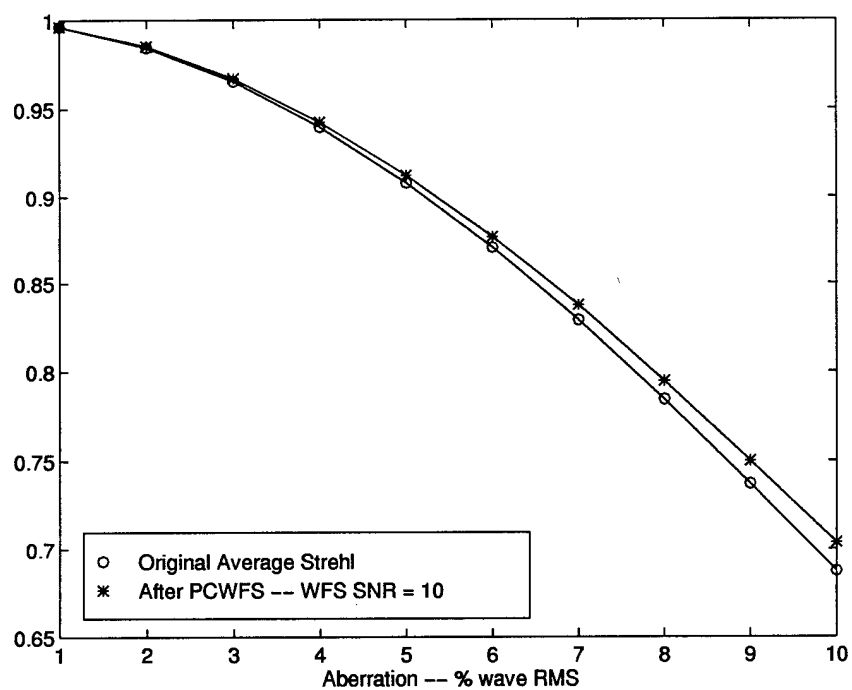


Figure 6. The Effect of OTF Phase Correction on average Strehl ratio. Each datum represents the average of an ensemble of 200 (effective) Strehl ratios, each derived from simulated imaging pupils with with random Gaussian aberrating phase screens of the specified RMS aberration strength. The PCWFS processed images used WFS data with a WFS SNR of 10.

Table 1. The effects of PCWFS on Strehl Ratio. See text for definition of quantities listed

P_{orig}	SR_{orig}	SR_{PCWFS}	P_{eff}
1	0.996	0.996	0.977
2	0.984	0.985	1.944
3	0.966	0.967	2.904
4	0.940	0.942	3.862
5	0.908	0.912	4.815
6	0.871	0.876	5.772
7	0.829	0.838	6.702
8	0.784	0.795	7.630
9	0.737	0.750	8.530
10	0.688	0.705	9.403

So from the *effective* PCWFS Strehl ratio, SR , an *effective* RMS wavefront error P_{eff} can be derived by inverting Marechal's approximation:

$$(34) \quad P_{eff} = \frac{50}{\pi} \sqrt{-\ln(SR_{PCWFS})}.$$

Simulation results for this effective aberration strength measure, and the effective Strehl ratio after phase correction processing are summarized in table 1

By inverting Marechal's approximation for the effective, phase-corrected Strehl ratios, table 1 tells us, for example, that a pupil that originally had 10% of a wave (RMS) of aberration acts like a pupil with 9.4 % of a wave of aberration, after PCWFS reconstruction, or a 6% reduction in effective RMS aberration in terms of an image plane quantity.

The table also shows us that Strehl ratios, already near unity, are pushed further towards this asymptotic limit by up to about 3% by using PCWFS.

2.6 Conclusion

Although the focus of this dissertation is the phase diversity wavefront sensing technique, derived in the next chapter, this chapter has first dealt with more generic simulation studies of wavefront sensor supported image post processing, using an *unspecified* simulated wavefront sensor, with user-definable error performance. The simulation analysis presented in this chapter thus serves as a feasibility study, or first-order “proof-of-concept” for the overall idea of using WFS information on space telescope imagery. This study is then a precursor to eventually using the phase diversity wavefront sensing technique in a deconvolution scheme similar the type shown here.

Results show that even when relatively noisy wavefront sensor information is used on images with up to 0.10λ RMS of unspecified wavefront error, the SNRs of image spectra can be boosted by a factor of 1.5 and RMS spectrum phasor angles can be approximately cut in half, across a wide range of spatial frequencies, for the simple case of noise-free point-source imaging. In the image domain, average effective Strehl ratios, already near unity, are pushed further towards this asymptotic limit by up to approximately 3%. This is achieved by straightforward, non-iterative post-processing manipulation of the Fourier phase of the image data, using information from some sort of wavefront sensing methodology. Furthermore, the complications inherent to inverse filtering, or expensive, complicated adaptive optics compensation are avoided.

III. Phase diversity wavefront sensing: theoretical considerations

The previous chapter was concerned with establishing the value of the method using WFS information in the deconvolution of weakly aberrated images. This treatment was general in the sense that no specific WFS methodology was considered. For the remainder of the dissertation the attention is focused on a single WFS methodology. In this chapter the theory for the phase diversity technique of aberration sensing is presented. The derivation is presented in terms of non-linear least-squares minimization, after Gonsalves (26). A maximum-likelihood treatment can be found in ref. (62). After developing the basic aberration sensing procedure, we examine the uniqueness and object-independence properties of the phase diversity procedure. The remainder of the chapter explains and justifies the regularization technique used in this application.

3.1 Least-squares phase diversity wavefront sensing

The phase diversity wavefront sensing (PDWFS) technique is an *a posteriori* image-based wavefront sensing technique which was first presented for optical imaging by Gonsalves (26). PDWFS measurements consist of multiple images of the same scene, each formed via a slightly different pupil phase screen. For example, when two images are collected, as in the typical PDWFS implementation shown in figure 1 (seen in chapter 1), one image, the “conventional” image, is aberrated only by the phase screen which is to be estimated. The second image, referred to as the “diversity” image, is aberrated by a different phase screen that has been perturbed from the original, unknown aberration in a *known* manner, perhaps by defocus as in figure 1. This known pupil phase difference is referred to as the *phase diversity*,

hence the name of the technique. In general, N images can be collected:

$$(35) \quad \{i_n(\vec{x}; \vec{\alpha}_n)\}_{n=1}^N = \{i_1(\vec{x}; \vec{\alpha} + \vec{\Delta}_1), i_2(\vec{x}; \vec{\alpha} + \vec{\Delta}_2), \dots, i_N(\vec{x}; \vec{\alpha} + \vec{\Delta}_N)\},$$

where i_n represents the n th diversity image, which is a function of image plane spatial coordinates, \vec{x} , and pupil aberration, $\vec{\alpha}_n$. The $\vec{\alpha}_n$ vector represents a set of aberration coefficients for some appropriate pupil phase basis set. For example, if

$$(36) \quad \phi_n(Rr, \theta) = \sum_{j=4}^J a_{n,j} Z_j(r, \theta)$$

as in equation 10, then the Zernike coefficients α_4 through α_J are arranged in a $(J - 3)$ -element aberration vector $\vec{\alpha}_n$:

$$(37) \quad \vec{\alpha}_n = [\alpha_{n,4}, \alpha_{n,5}, \dots, \alpha_{n,J}]^T,$$

with the T superscript denoting the transpose operation. This vector is in turn the resultant of two vector sub-components of interest: $\vec{\alpha}_n = \vec{\alpha} + \vec{\Delta}_n$. The fixed aberration to be estimated is designated by the coefficient vector $\vec{\alpha}$, unsubscripted, and the *known* phase diversity aberration is symbolized by the coefficient vector $\vec{\Delta}_n$. For the typical implementation of PDWFS (62) shown in figure 1, for example, the diversity for the defocused image would correspond to the a non-zero coefficient for the fourth Zernike polynomial, which corresponds to defocus (58).

In general, N diversity images are collected, as implied in equation 35. But for the derivation which follows, without significant loss of generality, we consider a pair of images collected using the phase diversity setup in figure 1. Later, the Gonsalves cost function which results for the two image case can be easily generalized to accommodate the N -image data set given by equation 35. Let the conventional

and diverse image data set be written as the output of an isoplanatic model in the image domain, as in equation 1:

$$(38) \quad \begin{aligned} i_1(\vec{x}) &= o(\vec{x}) * h_1(\vec{x}; \vec{\alpha} + \vec{\Delta}_1) \\ i_2(\vec{x}) &= o(\vec{x}) * h_2(\vec{x}; \vec{\alpha} + \vec{\Delta}_2), \end{aligned}$$

where $o(\vec{x})$ is again the unknown object irradiance pattern, $h_1(\vec{x}; \vec{\alpha} + \vec{\Delta}_1)$ and $h_2(\vec{x}; \vec{\alpha} + \vec{\Delta}_2)$ are the conventional and diversity PSFs, respectively, and $i_1(\vec{x})$ and $i_2(\vec{x})$ are the conventional and diversity images, respectively. Assume, for illustration purposes, that the phase screen is decomposed into Zernike polynomials, as in equation 36. In such a case, the phase screen quantities above could then be represented by $(J - 3)$ -element vectors, whose first element is the 4th Zernike coefficient:

$$(39) \quad \vec{\Delta}_1 = [0, 0, \dots, 0]^T = \vec{0}, \text{ the zero vector}$$

$$(40) \quad \vec{\Delta}_2 = [\delta, 0, 0, \dots, 0]^T,$$

where the symbol δ is used to represent the fourth Zernike coefficient of the second image collected by the system shown in figure 1, corresponding to the known defocus distance.

If the noisy data corresponding to $i_1(\vec{x})$ and $i_2(\vec{x})$ are denoted as $d_1(\vec{x})$ and $d_2(\vec{x})$ then

$$(41) \quad \begin{aligned} d_1(\vec{x}) &= o(\vec{x}) * h_1(\vec{x}; \vec{\alpha}) + n_1(\vec{x}) \\ d_2(\vec{x}) &= o(\vec{x}) * h_2(\vec{x}; \vec{\alpha} + \vec{\Delta}_2) + n_2(\vec{x}), \end{aligned}$$

where n_1 and n_2 represent the respective errors between the model and the data. Reference to the zero-valued $\vec{\Delta}_1$ diversity vector has been dropped. The integrated, squared value of the error, $(n_1^2 + n_2^2)$, is the quantity which is to be minimized under the least squares formalism:

$$(42) \quad J(\vec{\alpha}) = \int_{\vec{x}} \left\{ |d_1 - o * h_1(\vec{\alpha})|^2 + |d_2 - o * h_2(\vec{\alpha} + \vec{\Delta}_2)|^2 \right\} d\vec{x},$$

where, for simplicity, the spatial dependence (\vec{x}) of all d , o , and h quantities is now implied. The quantity $J(\vec{\alpha})$ is referred to as the phase diversity objective function or cost function.

At this point Parseval's theorem (22) and the Fourier transform convolution theorem are invoked in order to write

$$(43) \quad \begin{aligned} J(\vec{\alpha}) &= \int_{\vec{f}} \left\{ |D_1 - OH_1(\vec{\alpha})|^2 + |D_2 - OH_2(\vec{\alpha} + \vec{\Delta}_2)|^2 \right\} d\vec{f} \\ &= \int_{\vec{f}} \{N^2\} d\vec{f}. \end{aligned}$$

Notice that the \vec{f} spatial frequency dependence for all the D , O , and H quantities has been suppressed for simplicity of presentation. The symbol N^2 is used to denote the integrand of J , the total sum-amplitude-squared of the noise at some spatial frequency: $N^2 = N_1^2 + N_2^2$, where N_1 and N_2 are the Fourier transforms of n_1 and n_2 . Following (26, 43), the derivative of the integrand of $J(\vec{\alpha})$, with respect to the object quantity O is found at any given spatial frequency, and this derivative is set equal to zero,

$$(44) \quad \frac{\partial N^2}{\partial O} = -2(D_1 H_1^*(\vec{\alpha}) + D_2 H_2^*(\vec{\alpha} + \vec{\Delta}_2)) + 2O(|H_1(\vec{\alpha})|^2 + |H_2(\vec{\alpha} + \vec{\Delta}_2)|^2) = 0,$$

where the asterisk superscript represents complex conjugation. In writing 44, it is important to recall that the quantities being analyzed are complex in nature. This fact is accounted for by invoking the complex partial derivative identities shown in equations 2.69 and 2.70 of ref. (42). Solving for the least-squares object estimate \tilde{O} yields the expression for the Wiener-Helstrom filtered data (61),

$$(45) \quad \tilde{O} = \frac{D_1 H_1^*(\vec{\alpha}) + D_2 H_2^*(\vec{\alpha} + \vec{\Delta}_2)}{|H_1(\vec{\alpha})|^2 + |H_2(\vec{\alpha} + \vec{\Delta}_2)|^2}.$$

Inspection of equation 44 reveals that a second partial derivative of with respect to O would yield a positive, real number, guaranteeing that \tilde{O} minimizes N^2 at all spatial frequencies, thereby minimizing J . To obtain a suitable least-squares cost function for the inverse problem of estimating $\vec{\alpha}$ from D_1 and D_2 , \tilde{O} above is substituted back into $J(\vec{\alpha})$, yielding the Gonsalves phase diversity objective function, which after simplification is written as

$$(46) \quad J(\vec{\alpha}) = \int_{\vec{f}} \left(|D_1|^2 + |D_2|^2 - \frac{|D_1 H_1^*(\vec{\alpha}) + D_2 H_2^*(\vec{\alpha} + \vec{\Delta}_2)|^2}{|H_1(\vec{\alpha})|^2 + |H_2(\vec{\alpha} + \vec{\Delta}_2)|^2} \right) d\vec{f}.$$

The solution to the wavefront sensing problem now consists of finding the aberration vector $\vec{\alpha}$ which causes $J(\vec{\alpha})$ to reach its minimum value, or

$$(47) \quad \vec{\alpha}_{WFS} = \arg(\min_{\vec{\alpha}}(J(\vec{\alpha}))).$$

With this goal in mind, notice that the first two terms of the integrand of equation 46, the modulus-squared data, represent a constant for a given phase diversity data set. In other words, the quantity $(|D_1|^2 + |D_2|^2)$ is not a function of $\vec{\alpha}$ and need not be included for minimization purposes. Hence, the following simplified objective

function can be minimized

$$(48) \quad J(\vec{\alpha}) = - \int_{\vec{f}} \left(\frac{|D_1 H_1^*(\vec{\alpha}) + D_2 H_2^*(\vec{\alpha} + \vec{\Delta}_2)|^2}{|H_1(\vec{\alpha})|^2 + |H_2(\vec{\alpha} + \vec{\Delta}_2)|^2} \right) d\vec{f}.$$

This objective function was presented by Gonsalves in references (25, 26). It is not difficult to see that equation 48 takes the following form for the more general data set of N diversity images:

$$(49) \quad J(\vec{\alpha}) = - \int_{\vec{f}} \left(\frac{|\sum_{a=1}^N D_a H_a^*(\vec{\alpha} + \vec{\Delta}_a)|^2}{\sum_{b=1}^N |H_b(\vec{\alpha} + \vec{\Delta}_b)|^2} \right) d\vec{f}.$$

As noted in ref. (62), the same objective function can be derived as the negative of the image log-likelihood function, corresponding to a Gaussian noise model. A more appropriate, but much harder to use, Poisson model derivation of a maximum-likelihood objective function is also presented in ref (62).

Finally, note for the simulations used in the latter half of this chapter and in chapters 4 and 5 that all diversity phases will be due to defocus, introduced by adding known amounts of the fourth Zernike polynomial to the imaging pupil phase. With the exception of a single simulation case in chapter 5, all simulated phase diversity sets will consist of two images, one in-focus, conventional image and one defocused, diversity image. These two constraints will be relaxed in the novel application of phase diversity shown in chapter 6, where a more general interpretation of the overall technique is adopted—in chapter 6 the diversity phases are not confined to the family of parabolic phase screens, and the number of images is no longer restricted to 2.

3.2 Motivation for collecting diversity images

In this section two important properties of the Gonsalves cost function methodology are examined. First, an inspection of equation 49 shows that the objective function has no dependence on the original object $O(\vec{f})$ that is being imaged, which is a remarkable result. Second, the claim made previously regarding the uniqueness of PDWFS aberration estimates is discussed. These two properties stem from the fact that multiple, diverse images are collected, as opposed to a single image.

3.2.1 Object independence. In order to motivate the collection of two phase diverse images, compare the Gonsalves procedure with the corresponding single image problem. Assume that a single, in-focus image has been collected by an imaging system. In order to find the aberration that gave rise to that image, let us now attempt to parallel the least-squares development from the previous section.

Again, the goal is to minimize the modulus-squared error, this time for a single, in-focus image,

$$(50) \quad J(\vec{\alpha}) = \int_{\vec{f}} |D_1 - OH_1(\vec{\alpha})|^2 d\vec{f},$$

by finding the aberration vector $\vec{\alpha}$ which causes $J(\vec{\alpha})$ to reach its minimum value,

$$(51) \quad \vec{\alpha} = \arg(\min_{\vec{\alpha}}(J(\vec{\alpha}))).$$

These two expressions are the single-image version of equations 43, and a restatement of equation 47, respectively.

In the derivation of equation 49 it was shown that by finding the least-squares optimum object, it was possible to explicitly remove the object quantity from the

phase diversity cost function. In attempting to do the same for the case of a single image, the derivative of the integrand of $J(\vec{\alpha})$ can again be found, where the partial derivative is still with respect to the object quantity O , on a per-spatial-frequency basis. As before, this partial can be set equal to zero, and the least-squares object estimate, \tilde{O} , solved for. For the single image case, this yields the analagous, single-image \tilde{O} expression, which is given as (27, 43)

$$(52) \quad \tilde{O} = \frac{D_1 H_1^*(\vec{\alpha})}{|H_1(\vec{\alpha})|^2}.$$

This expression is the single-image version of equation 45, the Wiener-Helstrom or inverse-filtered data (65). Substituting this \tilde{O} back into $J(\vec{\alpha})$ yields

$$(53) \quad J_{min}(\vec{\alpha}) = \int_{\vec{f}} \left| D_1 - \left(\frac{D_1 H_1^*(\vec{\alpha})}{|H_1(\vec{\alpha})|^2} \right) H_1(\vec{\alpha}) \right|^2 d\vec{f}.$$

Inspection of this single-image version of equation 46 reveals a fatal flaw with this attempt at object independence for the cost function of a single image data set:

$$(54) \quad \begin{aligned} J_{min}(\vec{\alpha}) &= \int_{\vec{f}} \left| D_1 - \left(\frac{D_1 H_1(\vec{\alpha}) H_1^*(\vec{\alpha})}{H_1(\vec{\alpha}) H_1^*(\vec{\alpha})} \right) \right|^2 d\vec{f} \\ &= \int_{\vec{f}} |D_1 - D_1|^2 d\vec{f} \\ &= 0, \end{aligned}$$

which is the trivial result. Clearly, any attempt to explicitly solve for the object and thereby remove object dependence—as for the Gonsalves cost function of multiple, diverse images—is destined to failure for the case of a single image. So, in order to solve the single-image aberration estimation problem, one must instead retreat back to equation 50, attempting to minimize that object-dependent expression using trial

values of both O and $\tilde{\alpha}$ simultaneously, which is a vast complication. This single-image, object-dependent problem is essentially the least-squares-penalized blind deconvolution problem as formulated in refs. (40, 46). Hence, it is seen that it is the collection of the diversity image that allows a mathematically tractable, object-independent error metric to be derived.

3.2.2 Uniqueness and ambiguity in image-based WFS. Another motivation for collecting two or more diverse images stems from the fact that the mathematical mapping from the set of all possible phase screens to the set of all possible PSFs is, in general, a many-to-one mapping (24), as depicted in figure 7. This ambiguity causes the inverse WFS problem to be underdetermined when attempting to estimate pupil phase from a single image (24, 46, 47, 75). Ideally, the addition of a second, phase diverse image mitigates this uniqueness problem through overdetermination. In other words, this auxiliary image data introduces additional “automatic” or physical constraints on the phase estimate.

By way of comparison, notice that this uniqueness difficulty is also naturally manifest in the equivalent, single-image blind deconvolution/phase retrieval problem. The body of literature in that field, such as refs. (46, 47, 75), show that the standard strategies for attacking this problem also generally involve imposition of constraints. One primary difference between these techniques and phase diversity is that these constraints are of a more “common sense” or *ad hoc*, knowledge-based form. Examples of blind deconvolution constraints include imposing image positivity, and function support limitations of both the image and the image spatial frequency spectrum.

The concept of uniqueness in the general phase retrieval problem is difficult to attack quantitatively or analytically (75), except in special cases. A quantitative

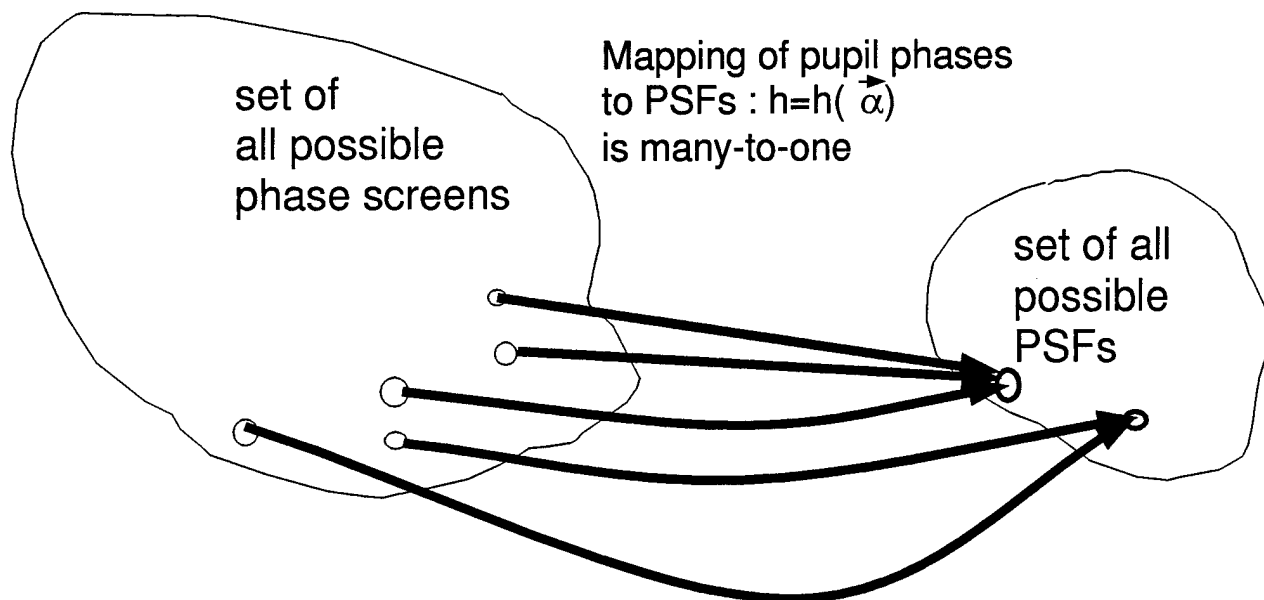


Figure 7. Conceptual diagram of the general phase retrieval uniqueness problem. Multiple phase screens can give rise to a single PSF/OTF.

investigation of uniqueness and ambiguity for the phase retrieval problem is shown in ref. (75). The research there showed, in a somewhat quantitative manner, the relative impact of knowledge-based constraints upon the phase retrieval problem for a large number of simple, random, simulated imaging cases. In this section, a single illustrative demonstration of the impact of phase diversity constraints is given.

The idea that multiple phase screens can give rise to the same PSF/OTF is easily demonstrated by the simple, computer generated point-source image shown in figure 8. In this figure the PSF shown in image (a) is due to phase screen (b), while the PSF shown in image (c) is due to phase screen (d). Phase screen (b) consists of 2.6 radians of Zernike mode 11, spherical aberration, while phase screen (d) exhibits -2.6 radians of the same aberration. Consequently, the two PSFs are identical to within the numerical roundoff error of the computer. The particular phase/image ambiguity exhibited here is a manifestation of a more general property

of PSFs which follows directly from Fourier optics. Specifically, a negation of the pupil phase causes a PSF to look as if it has been flipped about its spatial origin, or $h(\vec{x})$ becomes $h(-\vec{x})$. Of course, for a rotationally symmetric, spherically-aberrated example PSF, the flipped PSF looks exactly the same as the unflipped PSF.

The effect of this ambiguity on the single-image least-squares objective function of equation 50 is demonstrated in figure 9 (a). In this graph the single-image squared error metric for image (a) of figure 8 are plotted against trial values of the spherical aberration coefficient, which is the independent variable along the x-axis. In order to carry out this demonstration, the function being plotted, $J(\vec{\alpha})$ of equation 50, was modified to incorporate an *a priori* assumption of point-source imaging:

$$(55) \quad J(\vec{\alpha}) = \int_{\vec{x}} |d_1 - h_1(\vec{\alpha})|^2 d\vec{x}.$$

This simplified cost function is plotted versus trial values of the element of the $\vec{\alpha}$ corresponding to the 11th Zernike coefficient, with all other aberration coefficients set to zero. The quantity d_1 above is simply image (a) of figure 8. The ambiguity demonstrated qualitatively in figure 8 is now shown more quantitatively in this plot. Notice that the objective function plot (a) of figure 9 almost reaches a local minimum at $\alpha_{11} = -2.6$ radians of spherical aberration. This false local minimum almost matches the true local minimum at $\alpha_{11} = +2.6$, and in fact the two minima differ only by numerical round off error.

A similar plot can be created for the phase diversity scenario, by introducing a diversity image, d_2 , shown as image (e) of figure 8. Now the phase diversity version of equation 55 is written as:

$$(56) \quad J(\vec{\alpha}) = \int_{\vec{x}} \left\{ |d_1 - h_1(\vec{\alpha})|^2 + |d_2 - h_2(\vec{\alpha} + \vec{\Delta}_2)|^2 \right\} d\vec{x},$$

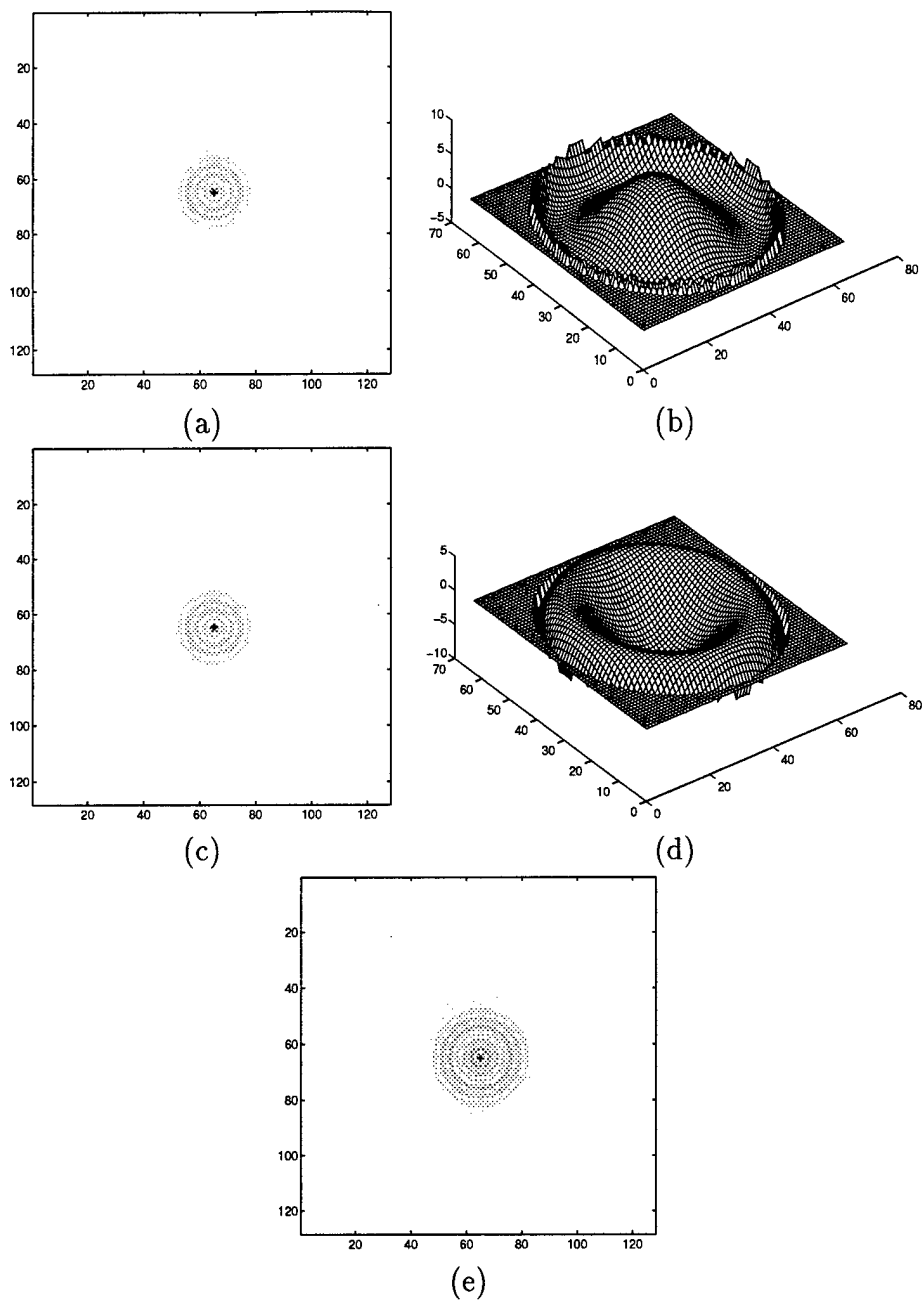
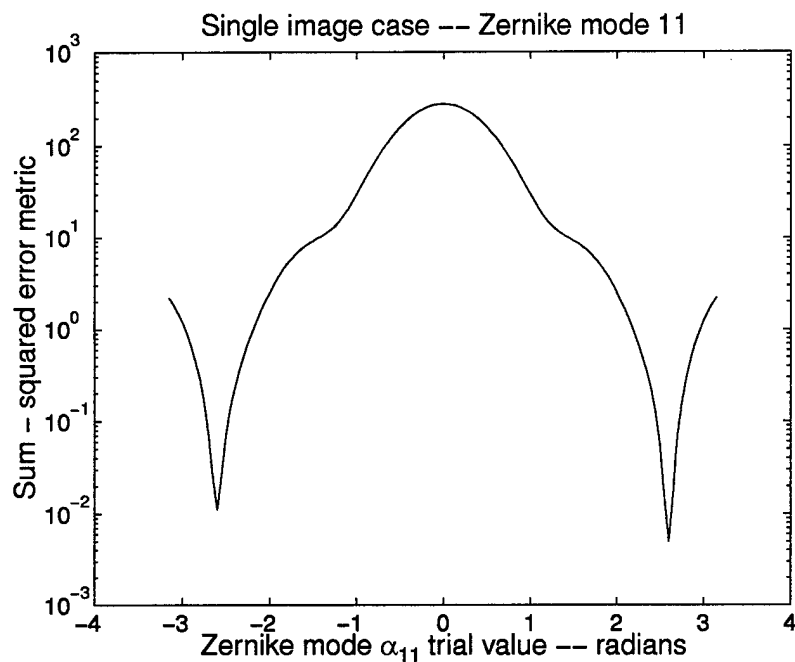
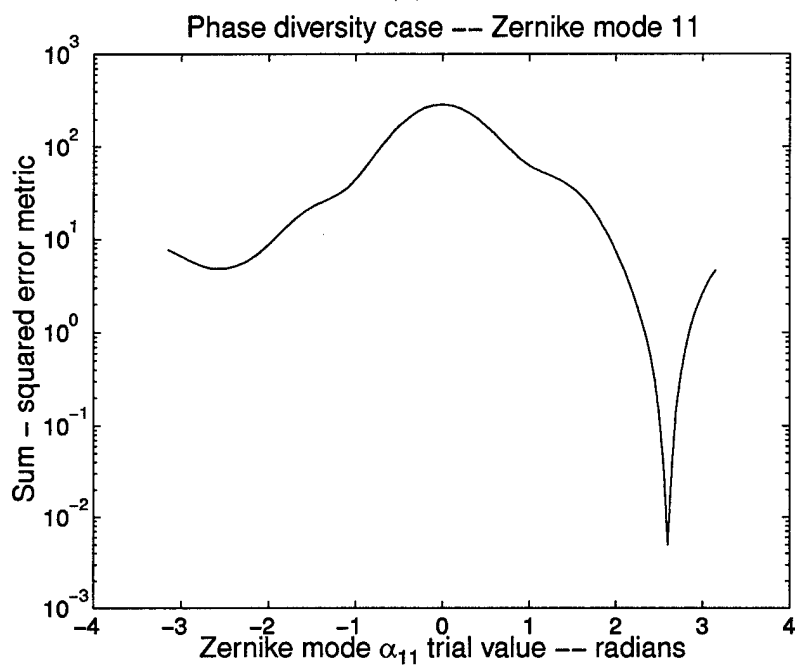


Figure 8. The conventional image (a) and original spherical aberration phase screen (b) used in the uniqueness demonstration described in the text. Image (c) is the ambiguous image which is due to the phase screen (d). Note that (d) is the negative of (b), yet (c) is identical to (a) to within round off error. Image (e) is the defocused, diversity version of image (a) used in the demonstration.



(a)



(b)

Figure 9. Single image (a) objective function (eq. 55), and phase diversity (b) objective function (eq. 56), for various trial values of Zernike mode 11, using the demonstration setup depicted in figure 8. Notice how the minimization ambiguity has been at least partially alleviated for this simple case by using the phase diversity approach.

again adopting, for demonstration purposes, the *a priori* assumption of point-source imaging. This error metric is again plotted against the spherical aberration coefficient in figure 9 (b). This plot shows that the ambiguity at $\alpha_{11} = -2.6$ has been partially alleviated by the introduction of a second, defocused image into the squared-error cost function, equation 56. Figure 9 (b) still reveals what appears to be a stationary point present at $\alpha_{11} = -2.6$, but the gradient of the cost function with respect to α_{11} is much less severe. This less-severe type of local minimum could be practically avoided through experimentation with various stopping criteria in the cost function minimizing routine, or even by executing multiple attempts at minimization, starting with a number of random initial guesses (30, 64).

3.3 Stabilization of the inverse problem

One final practical matter must be considered before minimizing equation 49, which is given again as:

$$(57) \quad J(\vec{\alpha}) = - \int_{\vec{f}} \left(\frac{|\sum_{a=1}^N D_a H_a^*(\vec{\alpha} + \vec{\Delta}_a)|^2}{\sum_{b=1}^N |H_b(\vec{\alpha} + \vec{\Delta}_b)|^2} \right) d\vec{f}.$$

Specifically, notice that

- quantities in the denominator may approach zero, and
- the recorded data are generally corrupted by noise.

These two effects can cause the minimization of the objective function to become unstable, or impossible to accomplish. Strategies for dealing with such instability are known as regularization (12).

3.3.1 Weighted least-squares noise suppression. In this work the strategy suggested in ref. (71) for regularizing the classic inverse filter problem will be followed.

Specifically, division by zero is prohibited in equation 57, with the integrand of $J(\vec{\alpha})$ set equal to zero at any spatial frequency where its denominator is zero. Also, the noisy data are filtered in the frequency domain with a “cone filter”, *i.e.* a circularly symmetric triangle filter with a user-selected cutoff frequency. The lowpass nature of the cone filter is a manifestation of the properties of photon noise, which dominates the higher spatial frequency content of a photon-limited image (29, 72).

If the cone filter is denoted by the symbol $F(\vec{f})$, and the set of all spatial frequencies within the union of the supports $H_1(\vec{f})$ and $H_2(\vec{f})$ is written as χ_1 , then the modified Gonsalves objective function becomes

$$(58) \quad J(\vec{\alpha}) = - \int_{\vec{f} \in \chi_1} \left(F \frac{|\sum_{a=1}^N D_a \tilde{H}_a^*(\vec{\alpha} + \vec{\Delta}_a)|^2}{\sum_{b=1}^N |\tilde{H}_b(\vec{\alpha} + \vec{\Delta}_b)|^2} \right) d\vec{f}.$$

A cone filter is shown in figure 10. The upper band limit of the cone filter is specified by a radial cutoff frequency ρ_{co} , where the symbol ρ denotes the Euclidean length of a spatial frequency vector \vec{f} . The equation of the cone filter is given, in terms of this radial frequency amplitude, as

$$(59) \quad F(\rho) = \begin{cases} 1 - \frac{\rho}{\rho_{co}} & \text{for } \rho \leq \rho_{co} \\ 0 & \text{otherwise,} \end{cases}$$

where ρ_{co} remains to be specified.

The cone-filtering technique can be put on firmer theoretical ground if it is interpreted as an implementation of the weighted least-squares variant of least-squares estimation theory. Under this interpretation, the cone filter is weighting the lower spatial frequency image information more heavily than the higher spatial frequency information. As stated in ref. (43), “[t]he rationale for introducing weighting factors

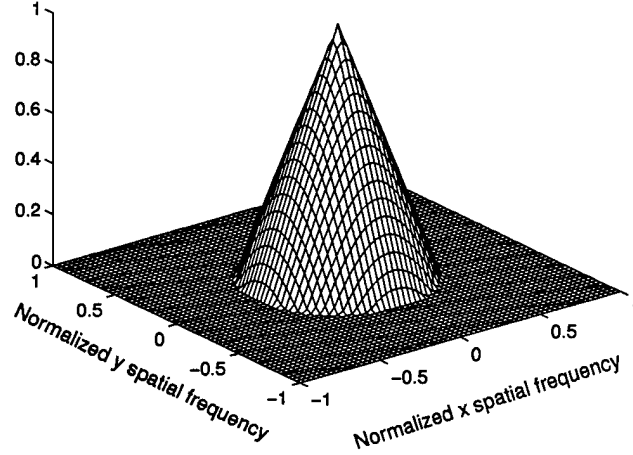


Figure 10. Mesh plot of a cone filter with some given cutoff frequency, showing the relative amplification performed by the filter in the spatial frequency domain.

into the error criterion is to emphasize the contributions of those data samples that are deemed to be more reliable (pg. 225)."

This reliability is quantified here in terms of sample-based estimates of the image spectrum signal-to-noise ratio (SNR) for the simulation-generated image ensembles used in the particular experiment. Recall from chapter two that the SNR of a spatial-frequency domain quantity $D(\vec{f})$ is defined as (29, 72):

$$(60) \quad \text{SNR}_D(\vec{f}) = \frac{|\langle D(\vec{f}) \rangle|}{\sqrt{\langle |D(\vec{f})|^2 \rangle - |\langle D(\vec{f}) \rangle|^2}},$$

where the angle brackets $\langle \dots \rangle$ indicate ensemble averaging. The SNR is seen to be the modulus of the mean image spectrum divided by the spectrum standard deviation. The cone-filter cutoff is chosen as that spatial frequency at which the radially averaged spectral SNR estimate drops below a value of 2. Implicit in this

choice is the assumption that frequency-domain information beyond this frequency limit is unreliable in a statistical sense (72). The low-pass nature of the cone filter means that the filter will act to exclude this noisy information at all higher spatial frequencies from the Gonsalves objective function minimization process.

Clearly the selection of cutoff frequency for this type of noise rejection filter involves a tradeoff between letting in too much noise and rejecting valid image information. Therefore, in actual implementation of the phase diversity technique, the selection of a regularization technique is a matter that deserves careful consideration. This straightforward regularization technique is used throughout all of the PDWFS estimation experiments presented in this dissertation. This simplification is imposed since a full analysis of regularization techniques for the phase diversity problem is beyond the scope of this work; the regularization topic could form the basis of a research project in itself.

There is support within the literature, however, for this simplified approach to noise suppression in inverse problems. For example, an SNR-based frequency domain weighting scheme is used in ref. (40) for filtering the least-squares penalty function of the penalized blind deconvolution technique — the single-image analog of the phase diversity problem. A data-driven approach to noise suppression filter design for PDWFS is given in ref. (49). Also note that the cone filtering stabilization approach can be interpreted as an implementation of the resolution kernel regularization method mentioned in ref. (76), where the inverse Fourier transform of the cone filter would represent the image-domain smoothing kernel referred to in that development. The idea of a noise-effective cutoff frequency in image reconstruction is also seen, for example, in ref. (72).

3.3.2 *Demonstration of regularization.* The utility of noise rejection filtering for Gonsalves phase diversity is demonstrated in figure 11. The plot shows coarsely sampled 1 dimensional slices of the Gonsalves objective function (equation 58) using noisy image data obtained from the simulation codes discussed previously. The two objective function curves are based on the same input image dataset; the only difference between the two situations is noise-suppression filtering. These two curves correspond to an objective function evaluation with the noise filter $F(\vec{f})$ as either

- a cone filter with a cutoff at 0.5 of the diffraction limit, or
- a “top-hat”, or unity-valued filter, which passes all spatial frequencies out to the diffraction limit.

The latter case corresponds to no noise suppression at all.

For this demonstration, the aberrated wavefront phase was characterized by a random $\lambda/4$ RMS aberration of Zernike modes 4 through 36. Trial values of α_7 , the 7th Zernike mode coefficient, are allowed to vary along the x axis of the plot, while all other aberrations are held constant. The actual value of Zernike mode 7 is 0.10 radians. The images were simulated as being very dim, 200 photons each, corresponding to a relatively high level of photon noise. This noise was simulated through a Poisson random number generator. The Poisson nature of photon noise is discussed in chapter 5. A defocused diversity image, with similar noise properties, was created by adding 2 radians of Zernike polynomial Z_4 .

The example shows how unsuppressed photon noise can cause an objective function minimization procedure to converge on an incorrect aberration value. In this demonstration the correct value of Zernike α_7 is 0.10 radians, a value which minimizes the filtered Gonsalves objective function, to within the sampling accuracy

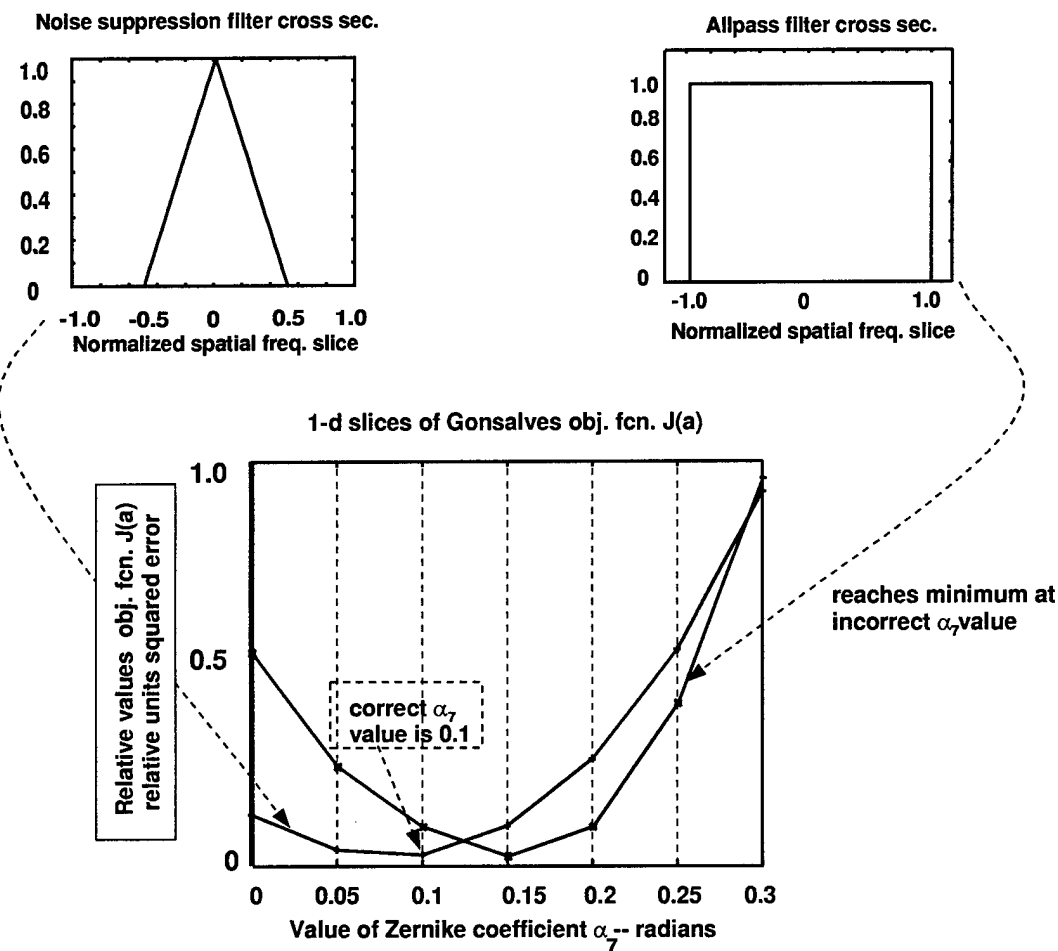


Figure 11. Demonstration of the effectiveness of the noise-rejection filter for the Gonsalves objective function. The plot shows a coarsely sampled 1 dimensional slice of the objective function both with and without a noise rejection filter. The original image suffered from approximately a quarter of a wave RMS Zernike aberration spread through modes 4 - 36. The actual value of Zernike mode 7, which varies along the x-axis of this plot, was 0.10 radians. The average number of photons was reduced to 200 per image for demonstration purposes. Without noise rejection, the objective function reaches a minimum at an incorrect point.

of this one-dimensional slice. On the other hand, the unfiltered objective function reaches its minimum at an incorrect value of $\alpha_7 \approx 0.15$ radians. This trend is borne out in numerous trial simulation runs used in the development of this research. If the data are not filtered to suppress noise, the phase diversity algorithm will naturally attempt to fit the WFS estimates to the noise, often resulting in inaccurate results. In fact, unsuppressed photon noise can prevent some iterative search routines from converging altogether. This failure to converge occurs because the unfiltered noise prevents the search algorithms from meeting reasonable stopping criteria. The random, noisy nature of the data can, for example, introduce false local minima into the objective function. These noise-induced minima can act to trap the search routine at an invalid phase estimate that is more due to the random image noise than the original aberrating phase screen.

3.4 Conclusion

This chapter has given the theoretical background for the phase diversity wavefront sensing (PDWFS) technique, which forms the focus of this dissertation research. Notice that the scope of our problem has been limited to least-squares estimation, under a linear, shift-invariant, photon-limited imaging model. Practical responses to the question “why bother with the collection of multiple, phase diverse images?” have been given in terms of the uniqueness and object independence introduced by the automatic constraints of diversity images. The issue of stabilization of the inverse problem by noise suppression filtering was also addressed.

IV. Monte-Carlo analysis of least-squares phase diversity: space telescope scenario

The goal of this chapter is to investigate the accuracy of the PDWFS technique in estimating the aberrations exhibited by a space telescope system. The investigation is carried out via numerical imaging simulations which model two different photon-limited space telescope imaging scenarios. The key assumption for the simulation models is that space telescopes are nearly diffraction-limited (16, 39, 41, 52, 50, 57). PDWFS performance is investigated via Monte-Carlo simulation experiments, a research effort that was suggested in ref. (61), but that has not been seen in subsequent published literature.

4.1 Introduction

Previous work by Paxman *et. al.* (60, 61) has demonstrated the use of the phase diversity technique for a sparse-aperture, six-element phased-array telescope whose images are corrupted by additive, white Gaussian noise. Demonstrations of phase diversity estimation of a total of 18 misalignment parameters are given in those references, along with an example of post-processing in the form of Wiener-Helstrom deconvolution.

This chapter expands on this previous work, moving from feasibility demonstration to Monte Carlo simulation of a number of different *photon-limited* space telescope imaging scenarios. Two different general categories of space telescope optical aberrations are modelled in the simulation. The first aberration case deals with Zernike modal aberrations, corresponding to the aberrations that might be experienced by a monolithic mirror that is vibrating or otherwise optically deformed.

The second case deals with segmented mirror misalignment aberrations for an approximate model of the proposed Next Generation Space Telescope (NGST) which features a 9-segment, mostly-filled aperture (1, 13, 14, 54, 79), as opposed to the sparse-aperture phased array analyzed in refs. (60, 61). Photon-limited (Poisson) point-source imaging is modelled, at three different light levels. The imaging pupils exhibit phase aberrations with overall average strengths of one-tenth of a wavelength pupil-averaged RMS, which is in the realm of "nearly diffraction-limited" imaging.

One of the most important aspects of this study is the inclusion of the effects of photon noise, an investigation which has not been addressed before in a Monte-Carlo sense. Key results from section 4.4 include an indication of a minimum acceptable light level for reliable PDWFS aberration sensing in the form of a noise-suppressed Gonsalves implementation. For point-source images averaging 1000 photons per image, the PDWFS estimation of NGST piston and tilt misalignment parameters proved unreliable in about 20 % of cases. Two cases with light levels at 10^4 and 10^5 photons per image did not exhibit this problem, and both yielded much more accurate WFS estimates. In fact, there appeared to be little difference between these two brighter cases in terms of WFS estimation performance. This aspect of PDWFS could have ramifications for the radiometric design of low-light imaging systems where photons are at a premium. Moreover, WFS accuracy also appears to be directly tied to the number of aberration parameters, or degrees-of-freedom (DOF), present in the problem. For instance, pupil-averaged RMS WFS errors nearly double when the NGST misalignment estimation problem is expanded to include segment tilt errors along with piston errors. The Zernike-aberrated case shows somewhat analogous WFS accuracy trends in terms of DOF.

Finally, a representative phase deconvolution example is shown, with results comparable to those of the generic WFS cases detailed in chapter 2.

The remainder of this chapter is organized as follows. Section 4.2 describes the two different types of aberrations that were modelled for these simulation experiments. Section 4.3 gives details of the computer simulation used in this chapter. The results of the simulations are presented in section 4.4.

4.2 Space telescope aberration bases

In this simulation effort two different pupil basis sets are used to create and specify space telescope aberrations. The statistical interrelation of these aberrations is discussed in section 4.3.1. One aberration model gives aberrations as linear combinations of the standard Zernike polynomials:

$$(61) \quad \phi_n(Rr, \theta) = \sum_{j=4}^J a_{n,j} Z_j(r, \theta),$$

as in the preceding two chapters. The underlying assumption is that the telescope system consists of a monolithic mirror which is experiencing aberrations that can be efficiently described by a relatively small number of Zernike modes. Such aberrations might, for example, be caused by spacecraft vibrations (11). The aberration vector is then given as the customary (J-3)-element vector

$$(62) \quad \vec{\alpha} = [\alpha_4, \alpha_5, \dots, \alpha_J]^T,$$

as in the derivation of chapter 3.

The second aberration case that we consider uses a zonal basis set corresponding to the pupil phase errors due to piston and tilt misalignments of the elements

of a segmented space telescope mirror. There has been some mention in the literature of preliminary plans for what is referred to as the Next Generation Space Telescope (NGST), a larger, more capable follow-on to the Hubble Space Telescope (HST). This set of simulation models is based on one possible configuration for this proposed 8 meter space telescope primary mirror. The most commonly proposed configuration generally involves 8 deployable "petals" surrounding an annular central mirror. Figure 12 shows a drawing of such a configuration that is under consideration by the National Aeronautics and Space Administration (NASA), as in refs. (1, 13, 14, 54, 79). Figure 13 shows the simplified approximation of the NGST pupil used here. Pupils were generated on a 64×64 discrete grid, with a pupil diameter of 64 pixels. The various optical features are highlighted in a gray-scaled map shown in figure 13 (a), with the piston error of the petals indexed sequentially so that they can be seen. A small, 6-pixel diameter central obscuration is included; this obscuration is artificially highlighted in figure 13 (a) to show its location. The mesh plot of figure 13 (b) shows a phase map corresponding to NGST pupil components that are randomly misaligned using the aberration basis set described next.

The 8 simulated petals are each allowed three degrees of freedom: piston and tip-tilt. The central annular mirror is allowed only two tilt degrees of freedom, making for an aberration vector $\vec{\alpha}$ of up to 26 elements. Notice that this means that any figure error within each individual mirror segment has been neglected. The

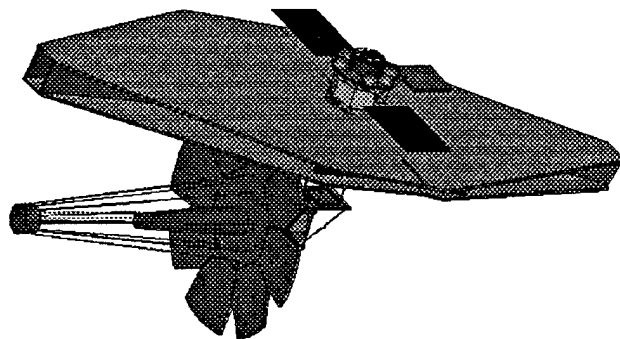


Figure 12. Possible configuration of the proposed Next Generation Space Telescope.

aberration vector for the NGST case, $\vec{\alpha}$, is arranged as follows:

$$\begin{aligned}
 \vec{\alpha}_{1-8} &= 8 \text{ petal piston errors} \\
 \vec{\alpha}_{9-16} &= 8 \text{ petal x-tilt errors} \\
 \vec{\alpha}_{17-24} &= 8 \text{ petal y-tilt errors} \\
 (63) \quad \vec{\alpha}_{25,26} &= 2 \text{ central mirror tilts, x and y.}
 \end{aligned}$$

The piston values represent the vertical phase displacement of the appropriate petal in radians. The various tilt parameters are proportional to the linear slope of a given segment in either one of two orthogonal directions. The phase map of figure 13 corresponds to some 26 random elements of the $\vec{\alpha}$ vector described above.

Note that, for telescopes with segmented primary mirrors, such as this NGST proposal, conventional slope sensors would be ill-suited to measuring the sudden pupil discontinuities introduced by misaligned segments (10). In this particular

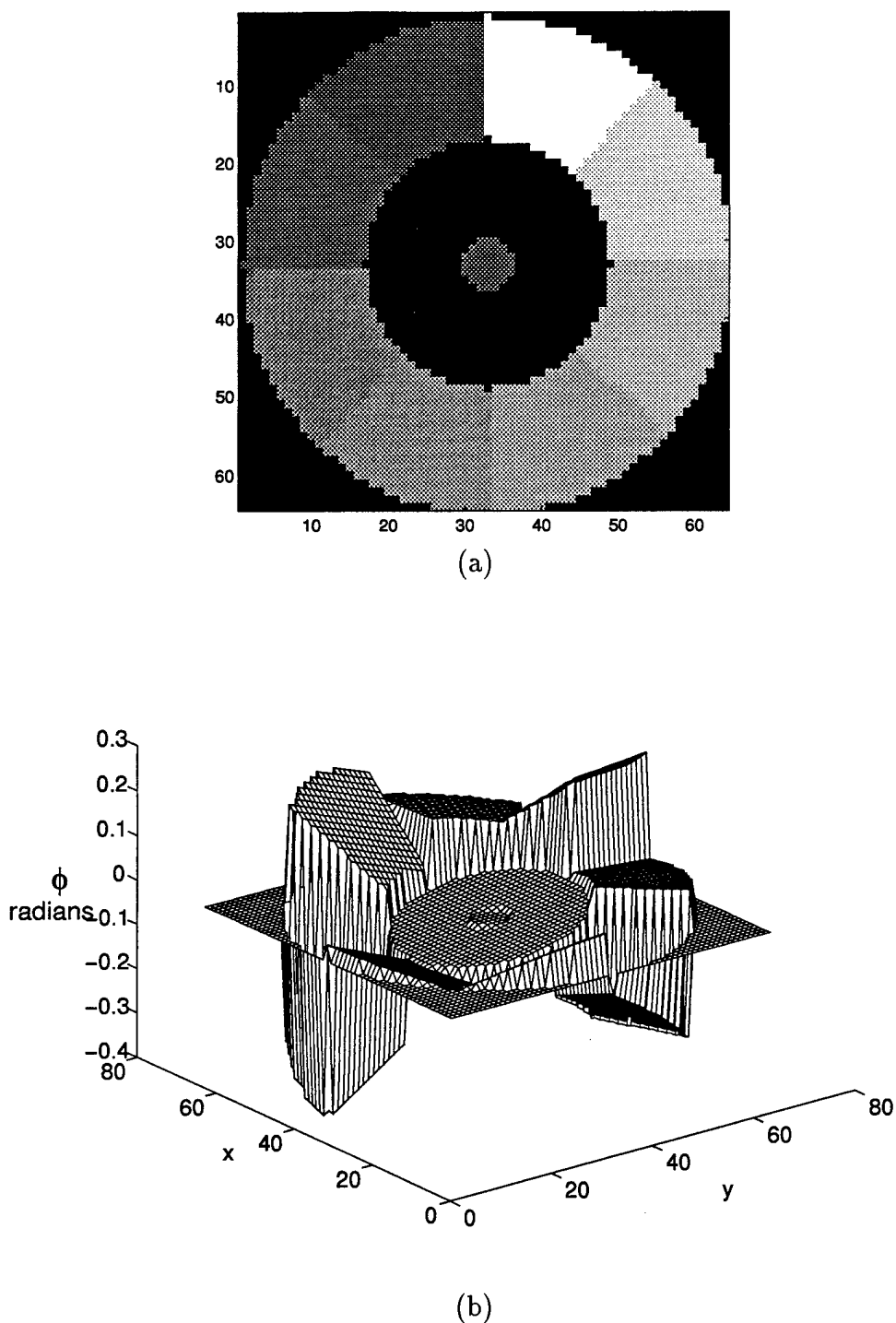


Figure 13. Simplified approximation of the pupil of the proposed 8-petalled Next Generation Space Telescope (NGST). A gray-scaled layout of the pupil components is shown in (a). A randomly misaligned NGST pupil phase map is shown in mesh surface format in figure (b). The pupils are sampled on a 64×64 discrete grid.

reference, the authors discuss how such segmented mirror section misalignments are measured in the ground-based Keck telescope using a prism and camera array which measures a point-source diffraction pattern at each segment discontinuity. Such a system is certainly not practical for an orbiting space telescope. PDWFS would present a simple, attractive alternative, provided the measurement accuracy is acceptable. This question is investigated by computer simulation.

4.3 Computer Simulation

A simplified flowchart of the numerical computer simulation for this chapter is shown in figure 14. This simulation incorporates the typical PDWFS configuration depicted in figure 1, with two images, where the second image is a defocused version of the first. In this section we discuss relevant details of the simulation model.

4.3.1 Random aberration generation. In the setup portion of the program the various simulation parameters are determined. The aberration case is selected, either Zernike or NGST. Then the spatial covariance matrix for the appropriate aberration coefficients must be specified. The numerical Monte Carlo imaging simulation procedure consists of generating circular pupil phase screens that are random linear combinations of the N desired aberration basis functions. Second order spatial statistics of the simulated phase screens are also specified and generated.

These second order statistics are specified and used as follows. For simplicity and generality of results, it was assumed that the various aberration coefficients are again zero-mean, statistically uncorrelated Gaussian random variables. In such a case, the standard method of generating correlated Gaussian random vectors, discussed in ref. (72), is utilized by first specifying an $N \times N$ covariance matrix, $[\Gamma_\alpha]$, for the aberration coefficient vector α_i . Since each aberration is simulated as being

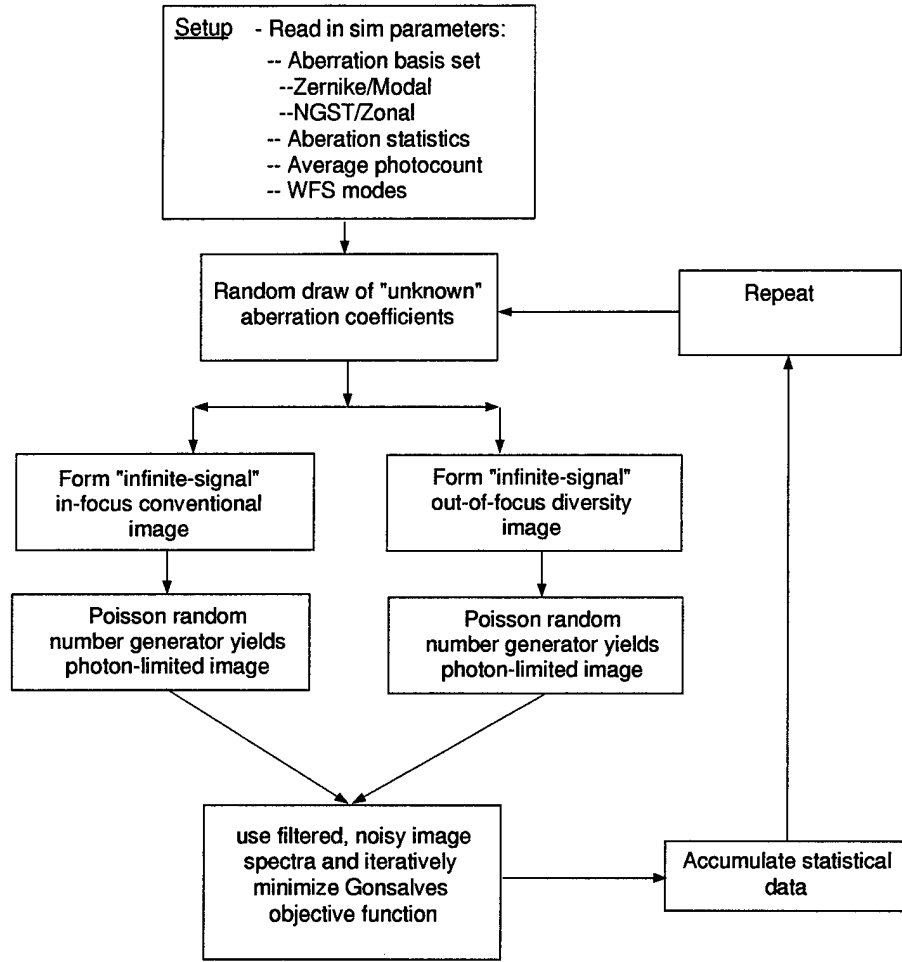


Figure 14. Computer simulation simplified block diagram.

uncorrelated with every other, the covariance matrices will always be diagonal. A realization of the zero-mean random vector $\vec{\alpha}$ is generated by the following matrix multiply operation,

$$(64) \quad \vec{\alpha} = R_{\alpha} \vec{b}$$

where \vec{b} is an N -element, uncorrelated, zero-mean, unit variance, Gaussian random vector. The matrix R_{α} represents the Cholesky factor, or in the diagonal case, simply the square root, of Γ_{α} .

This Monte Carlo generation of aberration vectors is carried out primarily in order to provide a thorough look at WFS performance with large numbers of combinations of different aberration coefficients. As in chapter 2, no attempt is made in this experiment, or in this entire dissertation, to rigorously model temporal or spatial aberration statistics for any particular space telescope. For instance, we can not reasonably expect the spatial aberration covariance matrix Γ_α to be a simple diagonal for a certain space telescope system. The statistical imaging characteristics of a given space telescope depend on the operational and mechanical aspects of the optics, spacecraft bus, and the overall space environment in which the telescope operates. However, even given all of these caveats, such a model is not unreasonable in light of our goal of simply exploring the aberration space in a Monte-Carlo sense. Additionally, it may still be reasonable, at least to first order, to accept the aberration generation scheme described here and in section 4.3.2 below as an approximate, simplified model of how spacecraft bus vibration and jitter are coupled into the aberration basis functions, similar to refs. (11, 57).

4.3.2 Simulation parameters. Two different general aberration cases were simulated for this chapter, each with two roughly analogous sub-cases:

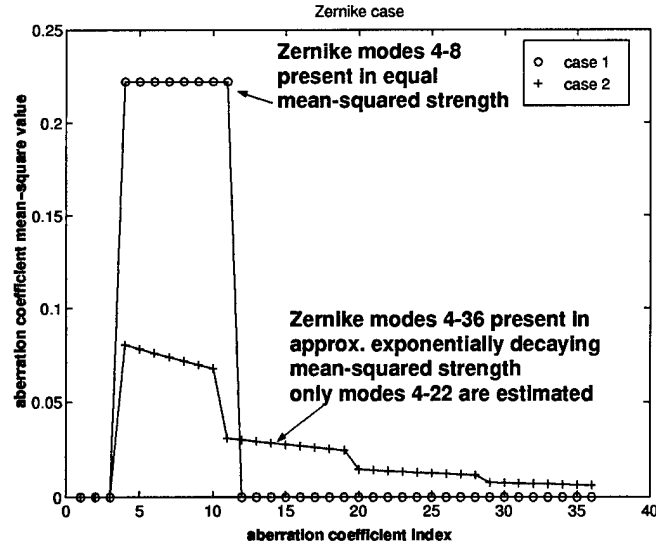
- Case 1a: 8 Zernike aberrations, modes 4-11, present in equal RMS strength. All 8 modes estimated by phase diversity.
- Case 1b: 8 NGST petal pistons present, all 8 estimated by phase diversity.
- Case 2a: Zernike modes 4-36 are present, with the average RMS strength of each mode order decaying in an approximately exponential fashion as mode order increases, as might be expected for a realistic monolithic mirror (11). Only modes 4-22 are estimated by phase diversity.

- Case 2b: All 26 NGST aberrations present in roughly equal proportion. All 26 aberrations are estimated by phase diversity.

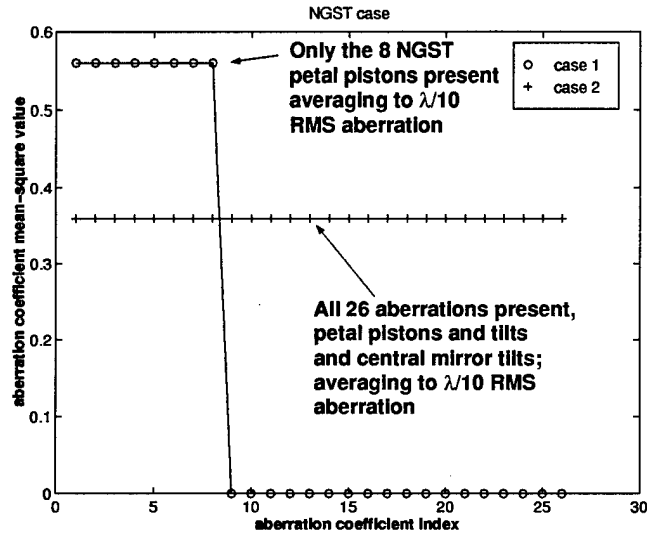
Cases 1a and 1b represent “easier” phase-diversity estimation problems in that a relatively small number of degrees-of-freedom (DOF) are present and estimated. These cases are roughly analogous to the piston-only cases of refs. (60, 61). Cases 2a and 2b represent more challenging and perhaps more typical phase diversity implementations, where a larger number of aberration DOF are present and estimated.

Figure 15 shows a plot of the diagonals of the 4 different aberration coefficient covariance matrices Γ_α , for the cases discussed above. These plots show the expected, ensemble average of the pupil-averaged mean-squared values of the randomly-drawn aberration coefficients, as the number of simulated phase screens tends to infinity. Figure 15 depicts the relative differences between the cases discussed above in terms of mean-squared aberration strength. Plot (a) represents the Zernike monolithic mirror modal aberration sub-cases, while plot (b) represents the NGST-like segmented mirror zonal aberration sub-cases. In all simulations the ensemble average of the pupil-averaged RMS wavefront error is $\lambda/10$.

As noted in section 4.1 these experiments incorporate photon-limited point source imaging, employing the semi-classical photodetection model (29, 72). Therefore, the quantity \bar{K} , the average number of photons present in each of the images, must also be specified in the simulation setup routine. The simulation model numerically incorporates a 50-50 beamsplitter assumption, with \bar{K} , the average image photon count, set the same for both the conventional and diversity images. Each of the experimental cases above were implemented for \bar{K} values of 1000, 10,000, and 100,000 photons. The total average number of photons incident on the entire imaging pupil is therefore $2\bar{K}$. Experimental runs consisted of 50 random imaging



(a)



(b)

Figure 15. Simulated aberration parameters—plots of the diagonals of the aberration coefficient covariance matrices Γ_α for the various experimental cases discussed in the text. These plots show the mean-squared values of the various aberration coefficients as generated for the Monte Carlo simulation realizations. The Zernike sub-case is given in plot (a) and the NGST sub-case in plot (b). We can see graphically the average strengths of the various aberration types in relation to each other for various simulated cases.

realizations apiece for the low DOF case 1, and 100 realizations for the high DOF case 2. Three light levels for four different sub-cases gives a total of 12 different phase diversity imaging scenarios.

Phase diversity estimation was accomplished via the modified Gonsalves algorithm discussed in section 3.3. Diverse images were perturbed by 2 radians times Zernike mode 4, (defocus), giving approximately the same quadratic diversity phase used in ref. (25). This choice of defocus parameter is discussed further in chapter 5. The regularized Gonsalves error metric was searched via a quasi-Newton method (30, 64, 66), starting with an initial aberration vector guess of the zero vector. Iteration stopping criteria accuracy limits on the aberration vector components and error objective function were all set as 0.001. Using the regularization filter design scheme of section 3.3, the cone filter cutoff for all cases was set at 50% of the diffraction-limited spatial frequency cutoff.

4.4 Simulation Experimental Results

This section shows the performance of least-squares phase diversity as a WFS technique in terms of pupil-averaged RMS error values. Results are presented in terms of individual realization and ensemble-averaged error values, and the estimation accuracy of individual Zernike modes for the appropriate cases. After reporting on the overall accuracy of the aberration estimation, the effectiveness of using these aberration estimates is demonstrated in a phase deconvolution example. In this chapter, the deconvolution technique is referred to as phase correction via phase diversity (PCPD), since we have now adopted a particular WFS technique.

4.4.1 Overall pupil-domain RMS WFS error. Figures 16 and 17 depict the pupil-averaged RMS wavefront sensing errors for the various cases simulated here,

on a realization-by-realization basis. The WFS error for the i 'th realization is given as:

$$(65) \quad \epsilon_{\text{wfs},i} = \sqrt{\left\{ \frac{\int d\vec{x} W(\vec{x}) (\phi_i(\vec{x}) - \tilde{\phi}_i(\vec{x}))^2}{\int d\vec{x} W(\vec{x})} \right\}},$$

where $\phi(\vec{x})$ is the actual truth-model pupil phase screen, $\tilde{\phi}$ is the estimated phase, and $W(\vec{x})$ represents the telescope aperture function (equation 6). The data points in these figures are sorted in ascending order according to ϵ_{wfs} value, for clarity of presentation.

The easiest general observation to be made from figures 16 and 17 is the spread of $\epsilon_{\text{wfs},i}$ values within any particular simulation case. For the ensembles simulated here, the WFS error values for a given case are distributed across approximately one-half to a full order of magnitude from lowest to highest. This sorted presentation shows, for example, that the technique can give noticeably different error performance when estimating two different wavefront phases of the same general statistical class.

One important performance limitation of the phase diversity WFS technique, as implemented here, is immediately revealed by examining plot (b) of figure 17. Experiments revealed that in the low light situation ($\bar{K} = 1000$) for case 2 b (the NGST 26 DOF case) the phase diversity search algorithm sometimes converged on a grossly incorrect aberration vector, with a phase estimation error several times larger than the actual $\lambda/10$ RMS original aberration. In this plot we can see that 20 of the 100 realizations exhibited a significant divergence of the phase diversity estimate from truth model phase screen values. These divergent cases indicate that when estimating a larger number of parameters, there is a certain threshold light

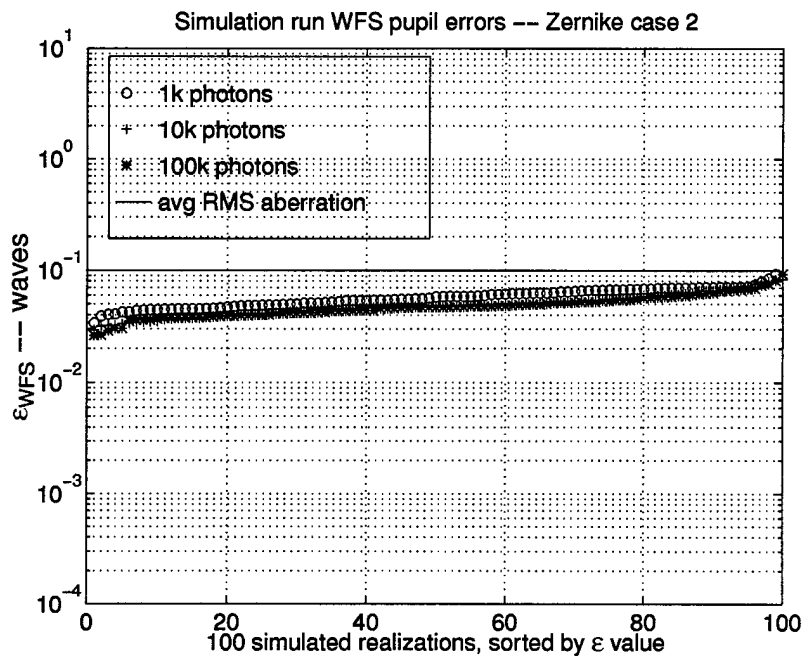
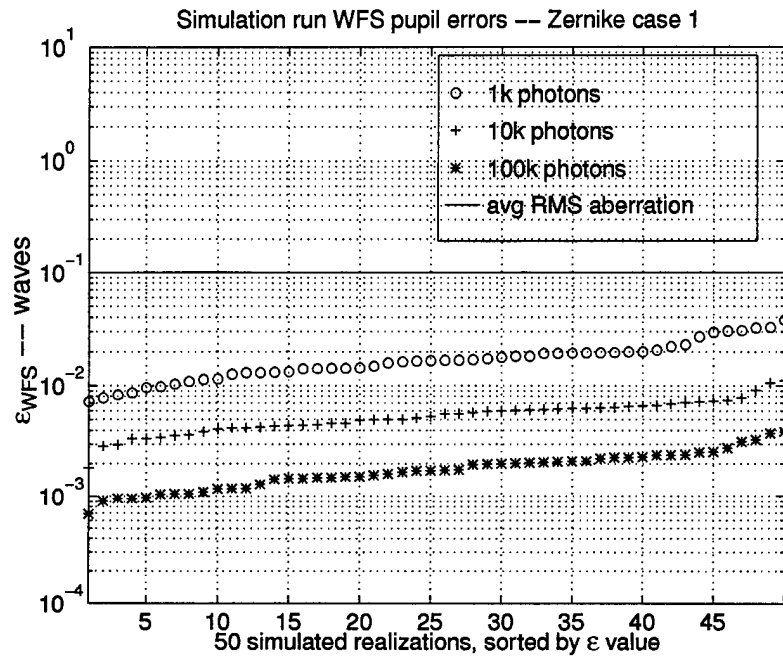


Figure 16. Pupil-averaged RMS phase diversity WFS estimation errors for the 50 simulated realizations of Zernike case 1 and the 100 realizations of Zernike case 2 (sorted by error value).

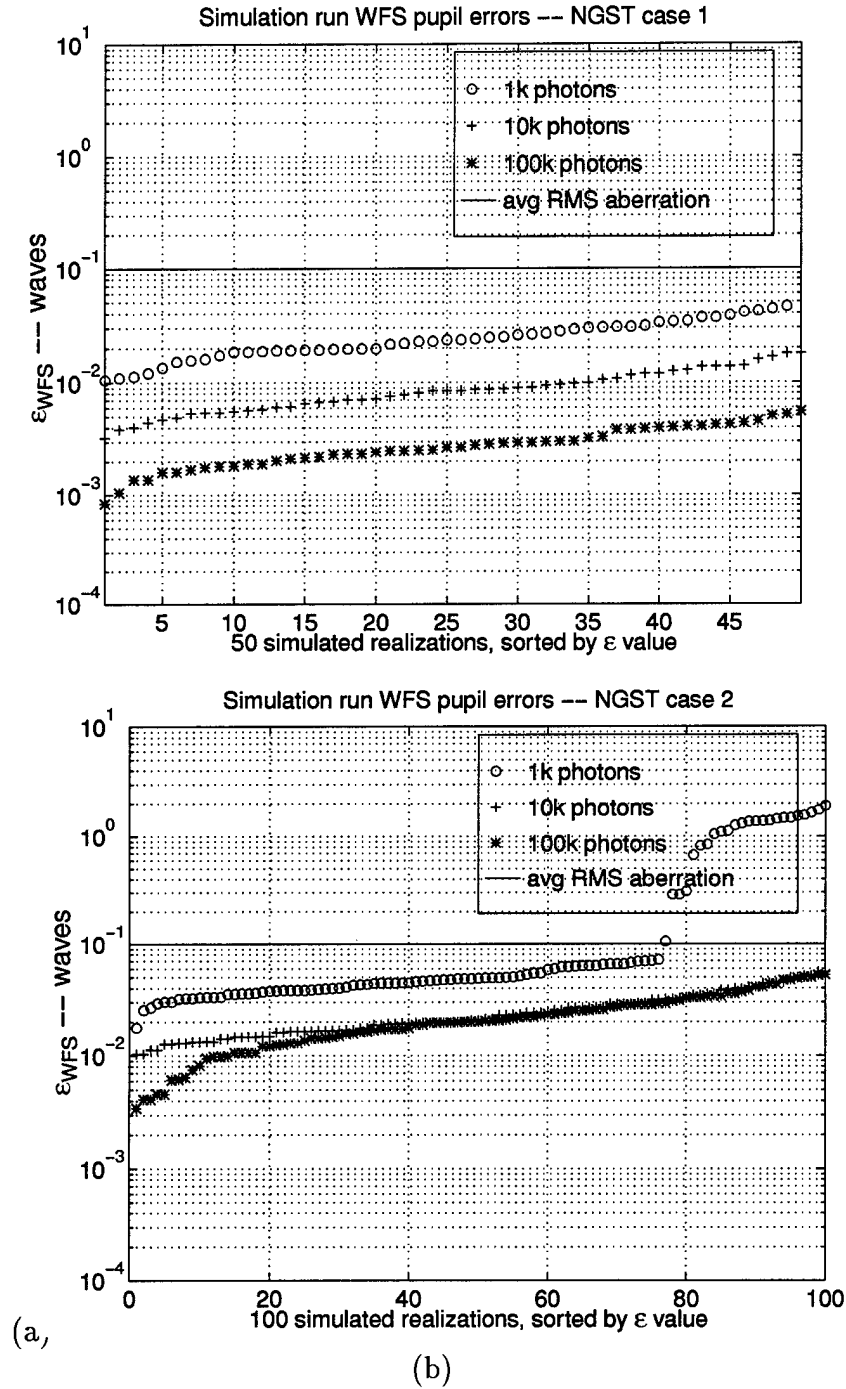


Figure 17. Pupil-averaged RMS phase diversity WFS estimation errors for the 50 simulated realizations of NGST case 1 and the 100 realizations of NGST case 2 (sorted by error value). Case 1 simulations exhibited and estimated 8 NGST petal piston errors. Case 2 simulations expanded upon case 1 by exhibiting and estimating segment tilt errors.

level below which photon noise can cause phase diversity aberration estimation to become unreliable.

4.4.1.1 Photon noise limitations of the least-squares PDWFS methodology. In order to discuss this result, let us note some facts about the Gonsalves technique. The least-squares formalism invoked in the section 3.1 is generally deterministic in nature — no probabilistic assumptions are made about the data. But, as shown in ref. (62), the Gonsalves objective function given in equation 49 can also be derived under the maximum-likelihood (ML) formalism, when it is assumed that the image data are Gaussian distributed random variables (43, 62) — in other words the Gonsalves objective function emerges as the negative of the Gaussian log-likelihood function.

Thus the Gonsalves technique can be cast as either as either a least-squares or a Gaussian-ML (GML) problem. Under either interpretation, the Gonsalves formulation has the advantage of simplicity: the optimum object estimate \tilde{O} can be solved for implicitly, and does not appear in the equation 49, as discussed at length in chapter 3. But the disadvantage of this formulation is that, as an ML problem, photodetection is generally not accurately modelled as a Gaussian random process, but rather a Poisson process. A Poisson-ML (PML) formulation of phase diversity also exists (62). The PML formulation has the advantage of correctly modelling photon noise, but, incidentally, has the significant drawback of depending explicitly on the object irradiance distribution, and thus has none of the simplicity of the GML/least-squares technique.

Finally, consider the well-known property of Poisson probability distributions where, in the limit of a large expected value (high light level), Poisson distributions (photon-limited images) asymptotically approach Gaussian distributions (44). This

statistical theorem leads to the conjecture that the Gonsalves and Poisson-ML techniques might give comparable performance when light is plentiful. This explanation is advanced in ref. (63) to explain experimental results that show a convergence of the Poisson-ML and Gonsalves (Gaussian-ML) techniques, when viewing the sun.

From the above considerations it is inferred that the GML and PML estimators are asymptotically equivalent, in the limit that there is enough light to cause photon noise to exhibit approximately Gaussian statistics. The results of plot (b) of figure 17, involving dim, point-source imaging, could possibly be indicating the approximate light level below which this GML approximation ceases to be acceptable. When imaging dimmer, and restricted or point-like objects—such as the simulated 26 DOF, 1000 photon NGST case under scrutiny here—the signal-dependent, Poissonian photon noise does *not* mimic the behavior of simple additive Gaussian noise.

Possible remedies for the low-light divergent behavior of least-squares PDWFS include:

- using different noise filtering technique (12, 49, 63);
- trying different amounts of defocus diversity (18), (chapter 5);
- repeating the minimization routine with a variety of different random initial guesses (64) and observing if several of them converge to the same answer; and finally
- utilizing the Poisson-ML technique (62).

4.4.1.2 Ensemble-averaged, pupil-averaged RMS wavefront estimation error. The average value of each of the curves shown in figures 16 and 17 is now found, in order to allow a more meaningful comparison of these WFS error results. Figures 18 and 19 depict the ensemble average of the pupil-averaged RMS wavefront

sensing errors for the various cases shown in figures 16 and 17, averaged across the simulation ensemble:

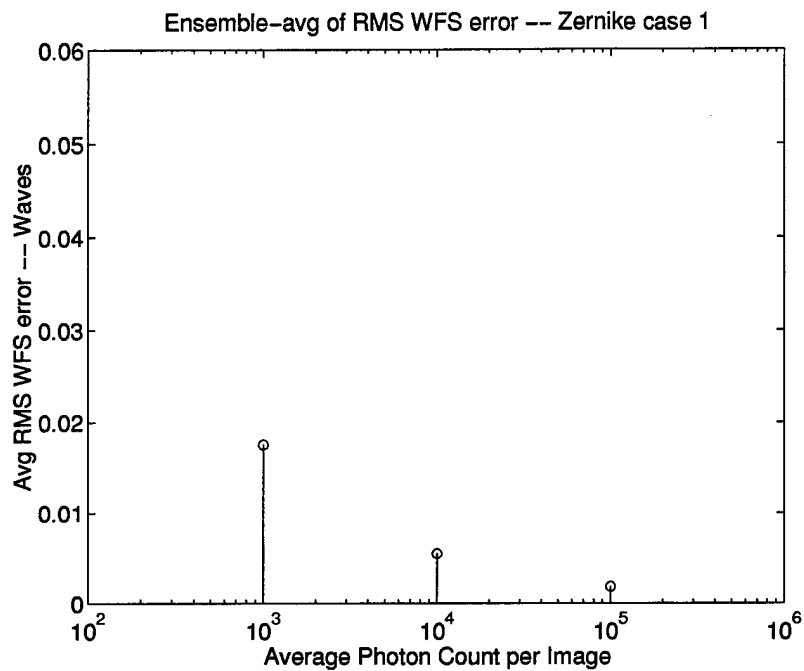
$$(66) \quad \langle \epsilon_{\text{wfs}} \rangle = \left(\frac{1}{N} \right) \sum_{i=1}^N \{ \epsilon_{\text{wfs},i} \}.$$

Notice for plot (b) of figure 19 that the 20 divergent WFS results discussed above have been excluded from averaging, and this data point should be compared to any others only provisionally.

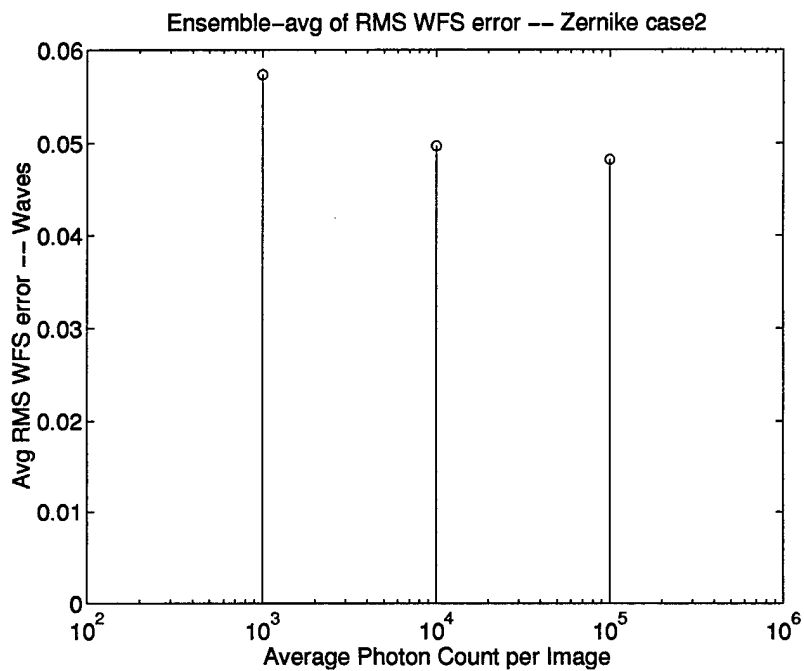
Figures 18 and 19 reveal that as light level decreases and/or the number of degrees-of-freedom (DOF) being estimated by phase diversity WFS increases, the overall trend is an increase in average RMS WFS estimation error. The (a) plots of these two figures show good WFS results as expected for the easier, low DOF estimation problems, and significant drops in error as light levels are increased.

Moving to the high DOF cases shown in the (b) plots of figures 18 and 19, we see a jump in the WFS error corresponding to the increased difficulty of the high DOF problem. In the low light case the jump from the low DOF problems to the high DOF problems results in either a doubling (NGST/zonal) or tripling (monolithic/Zernike) of the WFS error. We also note a certain sensitivity of the NGST/zonal problem to light levels by observing the significant decrease in WFS error as we increase the average number of photons by a factor of 10 from 10^3 to 10^4 average photons per image.

Another important observation regarding light level in the high DOF cases can be made by comparing the $\bar{K} = 10^4$ and 10^5 data points of figures 18 and 19. Once above what could arguably be called the dim “photon-starved” ($\bar{K} = 10^3$) realm, it does not seem to make a significant difference for the WFS technique when the light level is increased, even by an order of magnitude, from $\bar{K} = 10^4$ to 10^5 . This

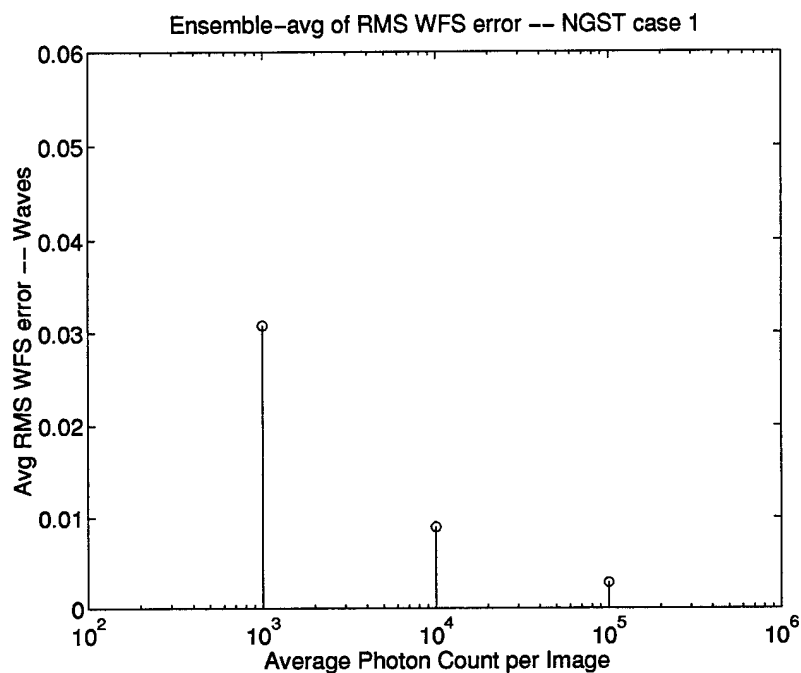


(a)

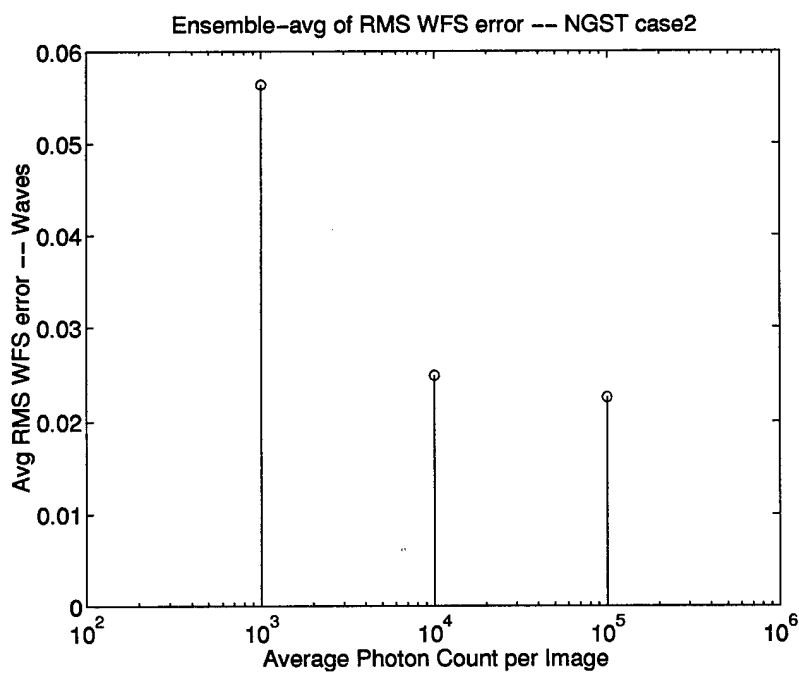


(b)

Figure 18. Averages of the curves in figure 16, giving ensemble averages of the pupil-averaged RMS estimation errors for (a) case 1 (b) case 2 of the Zernike aberration simulations.



(a)



(b)

Figure 19. Averages of the curves in figure 17, giving ensemble averages of the pupil-averaged RMS estimation errors for (a) case 1 (b) case 2 of the NGST aberration simulations. As noted in the text, the 20 divergent realizations for case 2b were excluded from averaging.

is a significant result in terms of system radiometry, when the technique is being considered for use in some low-light WFS application.

Notice that, even for the worst, low-light estimation cases shown in figures 18 and 19, there would still be a 40% reduction of the RMS pupil aberration if these estimates were used in adaptive optical mirror phasing scheme, as in ref. (10). This would represent a significant aberration correction, and corresponding boost in imaging capability, for a space telescope mirror that was originally aberrated by $\lambda/10$ RMS. When more light is available, the potential aberration correction would be even more substantial, as the WFS errors shown in figures 18 and 19 drop even further. The use of these phase estimates in post-processing, versus adaptive optical mirror phasing, is presented at the end of this chapter.

4.4.2 Zernike mode WFS errors. It should be clear from the layout depicted in figure 13 that any particular NGST petal is qualitatively indistinguishable from any other. In other words, it is clear that the petal basis functions differ from each other only in orientation. This means that the phase diversity WFS technique is on average no better in estimating the piston for petal 1 than it is for petal 5. The same can be said for NGST coefficients α_9 through α_{24} , the petal tilts of equation 63. On the average, no particular NGST pupil zone will be favored over another in terms of WFS estimation accuracy. In contrast, each of the Zernike polynomials are unique, and no such general characterizations about individual aberration mode sensing performance can be made.

In this section the distribution of WFS errors among the different Zernike modes is explored, as determined from the simulation. Specifically, the goal is to determine if there are any Zernike aberration modes that are estimated particularly well or particularly poorly by the phase diversity WFS technique.

Let us write the n 'th Zernike coefficient of the i 'th phase screen in an ensemble as $\alpha_{n,i}$. Then the ensemble-averaged RMS value of that coefficient is

$$(67) \quad (\alpha_n)_{rms} = \sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N \{\alpha_{n,i}\}^2}.$$

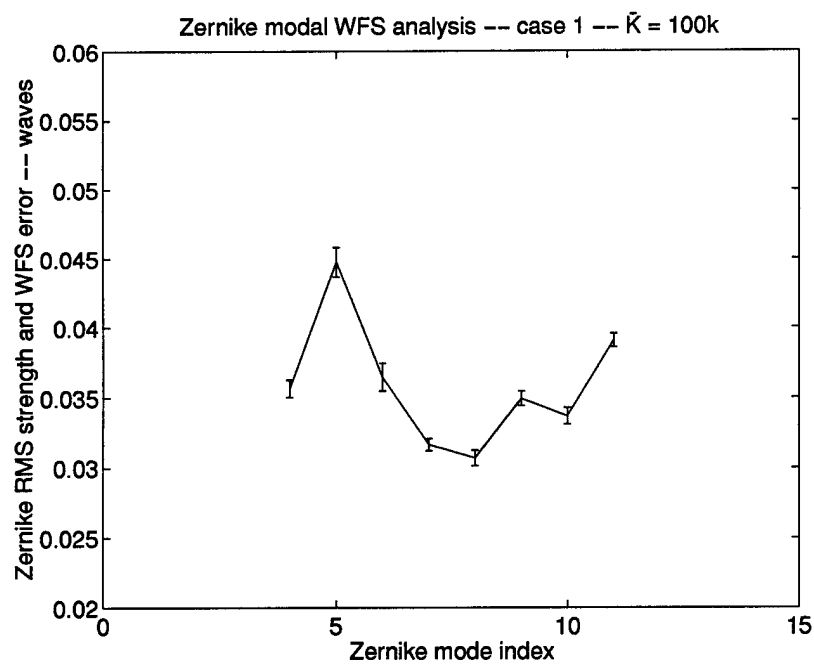
If the phase diversity estimate of $\alpha_{n,i}$ is denoted as $\tilde{\alpha}_{n,i}$, then the desired RMS mode estimation error can be written as

$$(68) \quad (\epsilon_{wfs,n})_{rms} = \sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N \{\tilde{\alpha}_{n,i} - \alpha_{n,i}\}^2}.$$

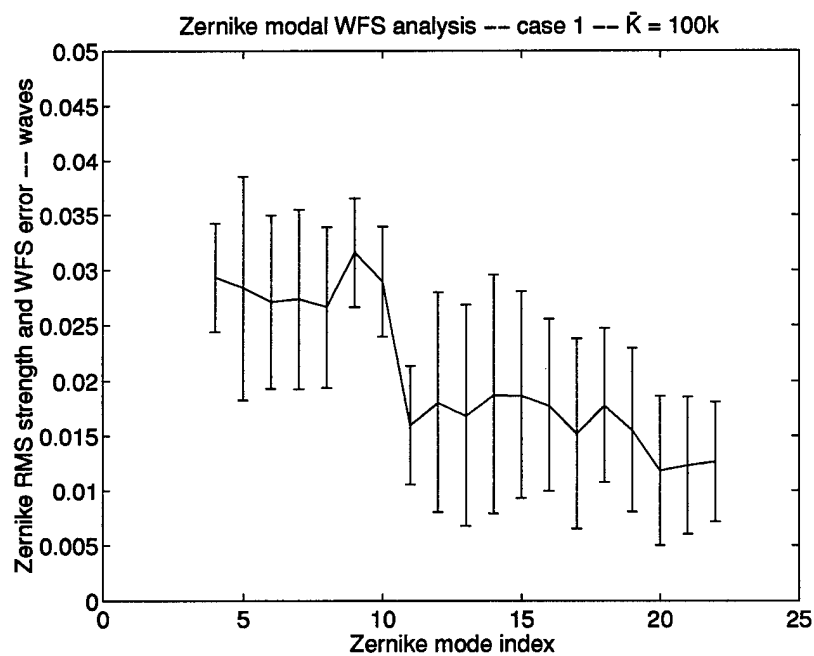
The solid line plotted quantities in figure 20 represent $(\alpha_n)_{rms}$, while the errorbar extent represents $(\epsilon_{wfs,n})_{rms}$ for the 8-mode and 19-mode cases. Only the $\bar{K} = 10^5$ cases are shown, with the other cases simply showing a correspondingly larger extent to the errorbars as expected from the data in figure 18. The WFS error seems to be approximately equally distributed across the Zernike modes in an RMS sense. These data support the hypothesis that there are no particular Zernike modes that are estimated any better or worse than any of the others, among the first 22 modes.

4.4.3 Using PDWFS estimates in phase deconvolution. In this section we will refer to the phase deconvolution technique studied in chapter 2 as phase correction via phase diversity (PCPD). The PCPD object estimator was given in equation 20, where the OTF estimate is used in a conjugate unit phasor, which acts to correct the Fourier phases of the detected image spectrum D . Following chapter 2, this section will show how the PCPD estimator performs given the phase diversity aberration estimates discussed above, specifically in terms of

- improving the image spectral SNR,



(a)



(b)

Figure 20. $\bar{K} = 10^5$, sample-based average Zernike aberration mode strength and WFS RMS error bars, for (a) case 1a, and (b) case 2a.

- reducing the image spectrum phase errors, and
- improving Strehl ratio.

PCPD Spectral SNR Improvement

The simulation model used here includes two sources of randomness in the original imaging system:

1. randomness of the pupil aberrations from one realization to the next; and
2. random photon arrival, as described by the semi-classical model of photo-detection.

Item 2 manifests itself both in the original image that is being deconvolved, and in the error present in the WFS estimate, as discussed in great detail in section 4.4.1.

The curves shown in figures 21 through 24 depict radially-averaged profiles of estimates of the quantity $SNR_H(\rho)$, the SNR of the effective optical transfer function before and after PCPD processing. The variable ρ , consistent with the radial-averaged nature of the plots, denotes the amplitude of the spatial frequency vector, and is normalized to 1 at the diffraction limit:

$$(69) \quad \rho = |\vec{f}| = \frac{D\sqrt{f_x^2 + f_y^2}}{\lambda},$$

where D is the physical telescope aperture diameter, and λ is the center imaging wavelength.

The SNR quantities are simulation sample-based estimates, each obtained by averaging over ensembles of computer generated imaging realizations of the various experimental cases. In order to condense the presentation of the results, only the low and high light level cases are shown ($\bar{K} = 10^3, 10^5$ per image) to illustrate the

trends. The SNR metric for the intermediate case of $\bar{K} = 10^4$ per image provides no new qualitative information that cannot be gained from the other two extremes.

The trend across all of the SNR plots (figures 21 through 24) clearly tracks the WFS performance trends of figures 18 and 19. Specifically, the degree of OTF SNR increase gained from PCPD seems to be clearly tied to the level of WFS estimation accuracy, as presented in the previous section. The more accurately the aberrated OTF is estimated, the greater the benefit obtained from using this information in frequency domain phase correction.

The degree to which photon noise corrupts higher spatial frequency information in the phase diversity estimation process is shown dramatically by comparing the (a) and (b) plots in any of figures 21 through 24. A much greater degree of resolution improvement is seen, for example, in the $\bar{K} = 10^5$ case of figure 21 (a), where the noise-effective cutoff (NECO) frequency, the frequency at which the SNR drops below 2, is extended from 0.8 normalized in the original image, out to 0.97 normalized after PCPD processing. Compare this to the $\bar{K} = 1000$ case of figure 21 (b), where the NECO is increased from 0.7 to 0.8 normalized. The PCPD SNRs fall off much more rapidly with spatial frequency for the low-light cases than for the cases with brightest light levels. The degree of SNR improvement corresponds with the level of WFS accuracy, as discussed in section 4.4.1.2. For the high \bar{K} , low DOF cases shown in figures 21 and 22, for instance, the highly accurate WFS estimates enable the PCPD algorithm to boost the NECO out to nearly the diffraction limit. These results also generally correspond with the results of the feasibility study for a generic wavefront sensor in chapter 2.

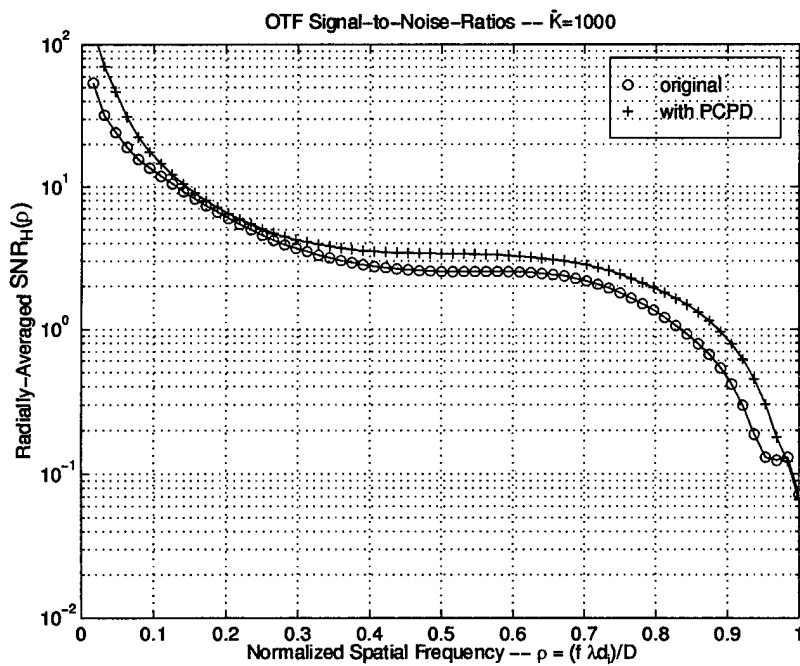
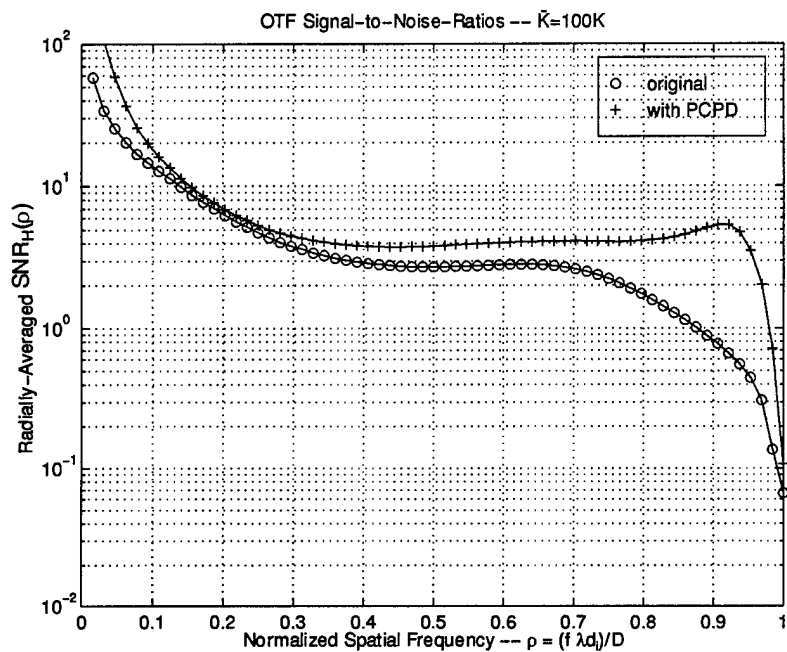


Figure 21. Case 1a — Zernike, low DOF. Radially-averaged profile of OTF Signal-to-Noise Ratios before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$; The expected values needed for each curve are estimated by averaging over the entire ensemble of simulated realizations.

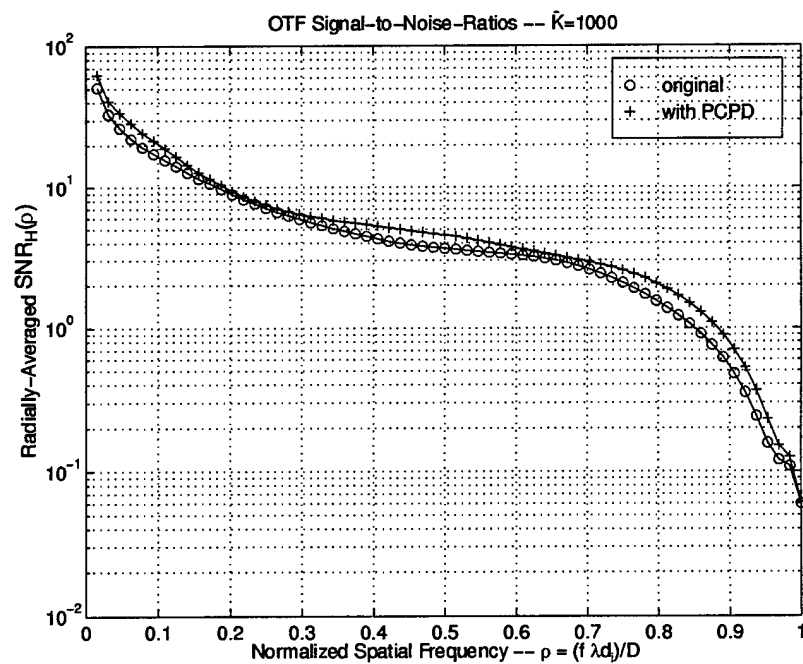
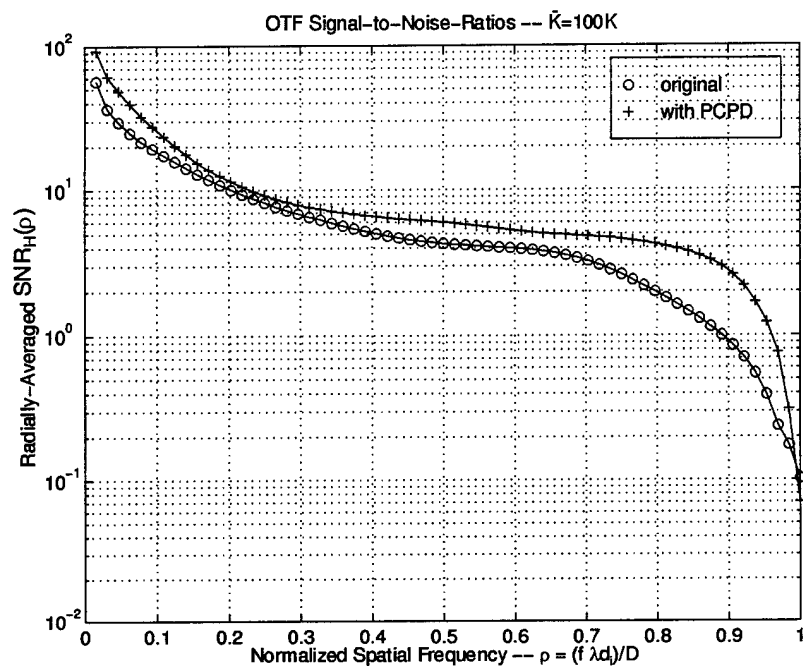
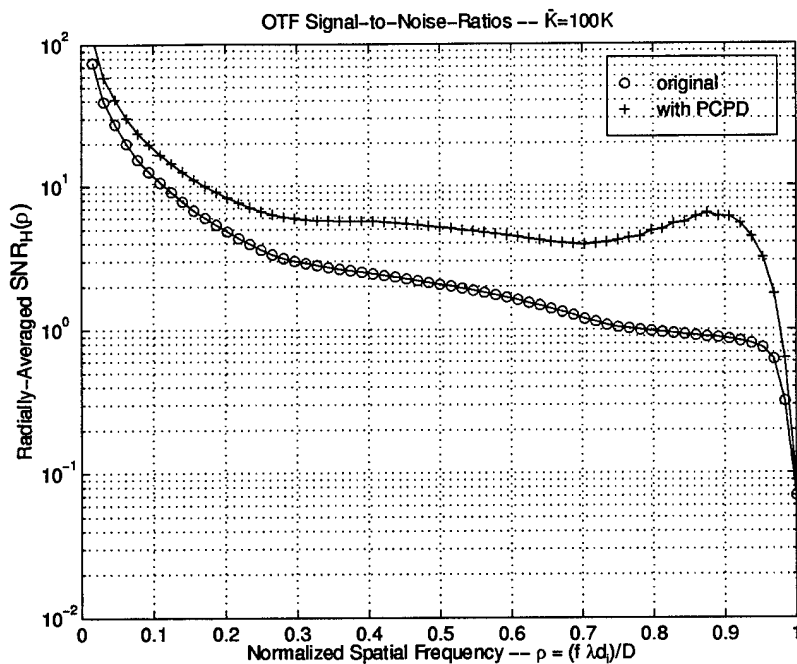
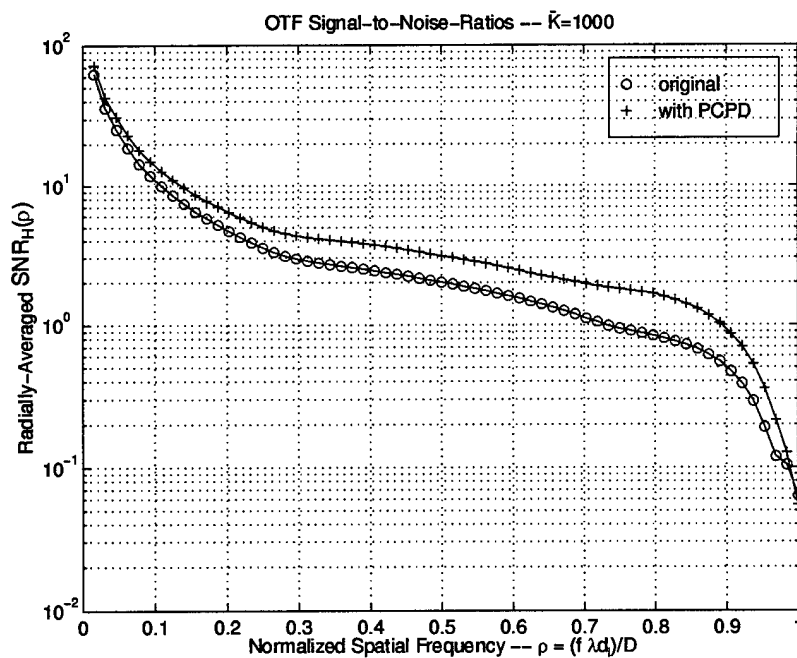


Figure 22. Case 2a — Zernike, high DOF. Radially-averaged profile of OTF Signal-to-Noise Ratios before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$; The expected values for each curve are estimated by averaging over the entire ensemble of simulated realizations.

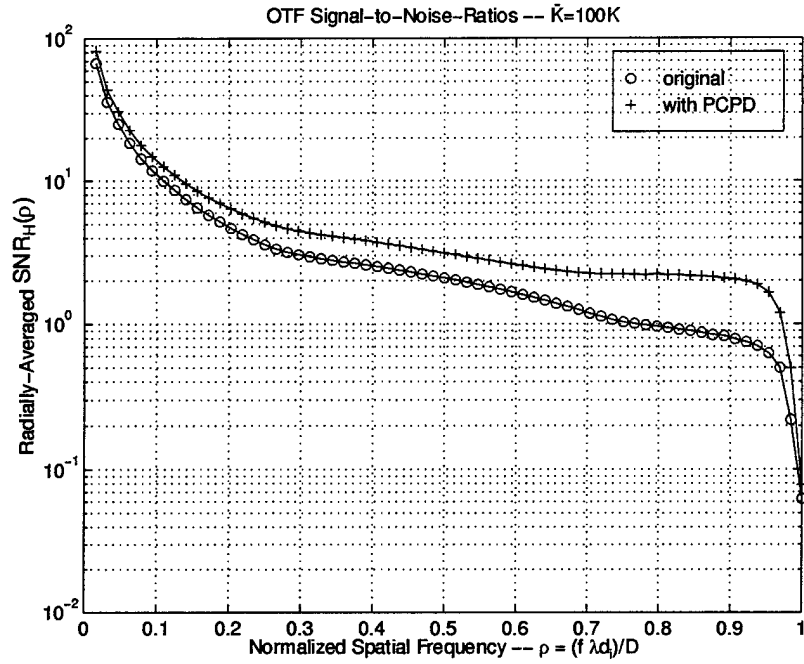


(a)

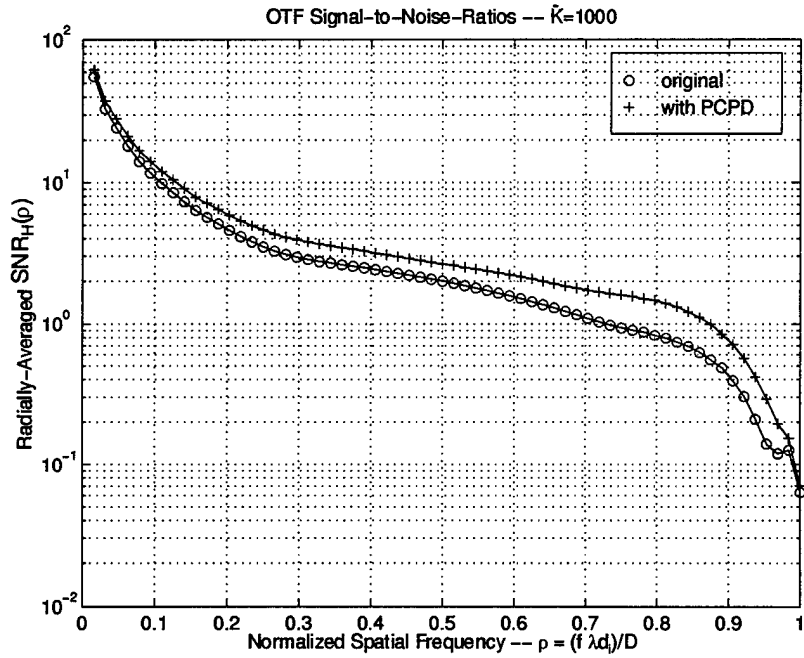


(b)

Figure 23. Case 1b — NGST, low DOF. Radially-averaged profile of OTF Signal-to-Noise Ratios before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$; The expected values for each curve are estimated by averaging over the entire ensemble of simulated realizations.



(a)



(b)

Figure 24. Case 2b — NGST, high DOF. Radially-averaged profile of OTF Signal-to-Noise Ratios before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$; The expected values for each curve are estimated by averaging over the entire ensemble of simulated realizations, except for (b), where the divergent WFS realizations are excluded, as discussed in the text.

PCPD Spectral Phase Error Reduction

Recall that the goal of PCPD is to drive the phasor angle of the effective OTF to zero (equation 20) in the limit of perfect OTF estimation by the particular WFS technique ($\epsilon_{wfs} \rightarrow 0$). Ideally, the effective OTF should be an entirely real-valued quantity, with the complex phasor angle of each spatial frequency component of the effective system OTF equal to zero. Figures 25 through 28 give a detailed picture of the RMS phasor angle behavior of the effective PCWFS OTF is at various spatial frequencies. Specifically, each of the curves represents a radially-averaged profile of the quantity $\Phi_{rms}(\rho)$, where

$$(70) \quad \Phi_{rms}(\rho) = \text{Radial Average} \left[\sqrt{\frac{1}{N} \sum_{i=1}^N [\Phi_{H_i}(\vec{f})]^2} \right],$$

and $\Phi_{H_i}(\vec{f})$ represents the phase angle of the particular OTF in question, at spatial frequency \vec{f} . Again, the ρ variable represents the normalized amplitude of the spatial frequency vector, the independent variable that remains after radial averaging of a frequency-domain quantity. Each curve is a radially-averaged profile of the RMS value of the complex phasor angle error introduced into the image by effective OTF.

All of these OTF phasor-based plots show quantitatively that, as expected, the values of the effective PCWFS OTFs are driven towards the real axis with varying degrees of effectiveness, again dependent upon the accuracy of the WFS estimation, as quantified in section 4.4.1, completely analogous to the results of chapter 2.

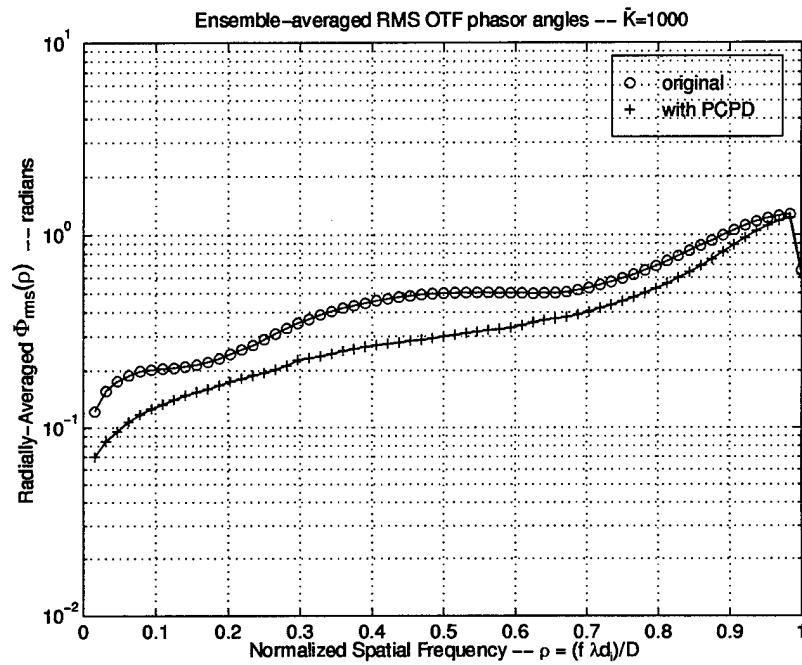
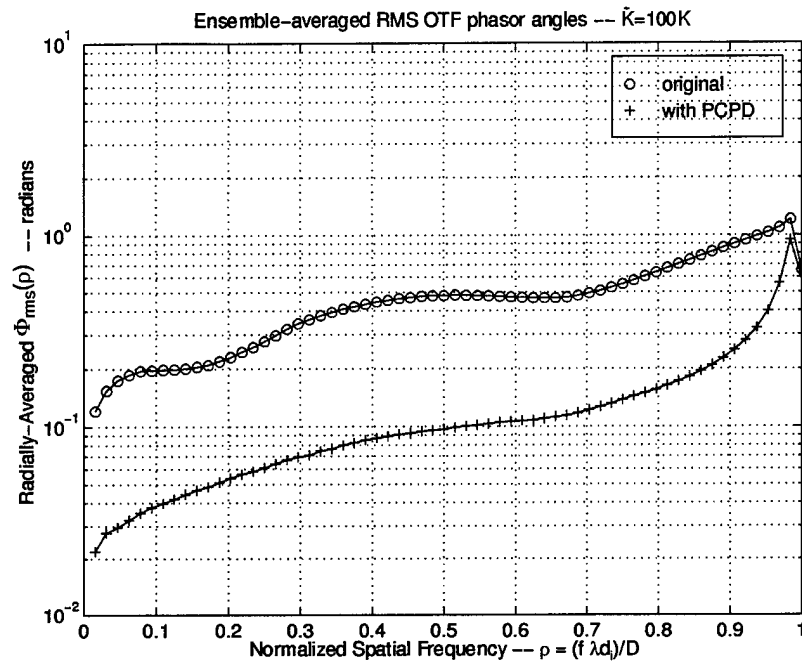
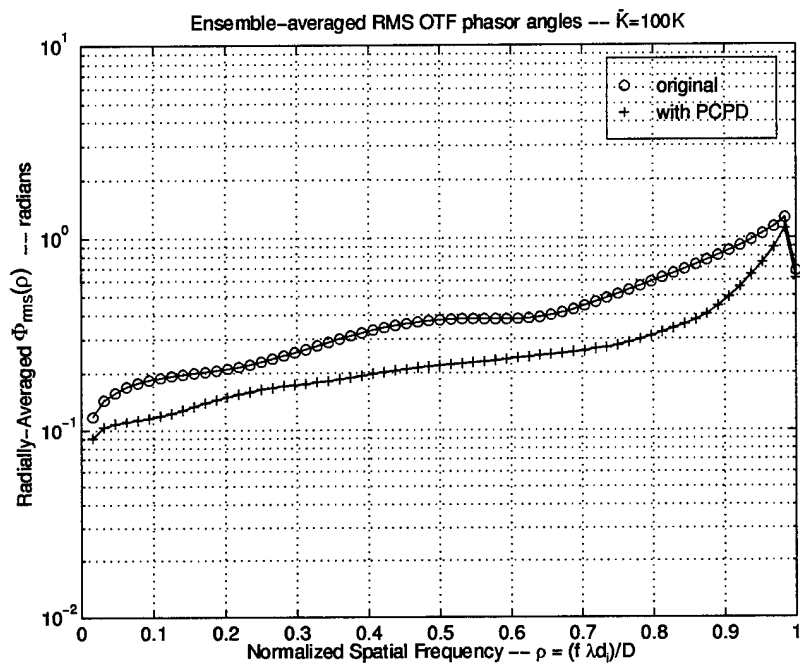
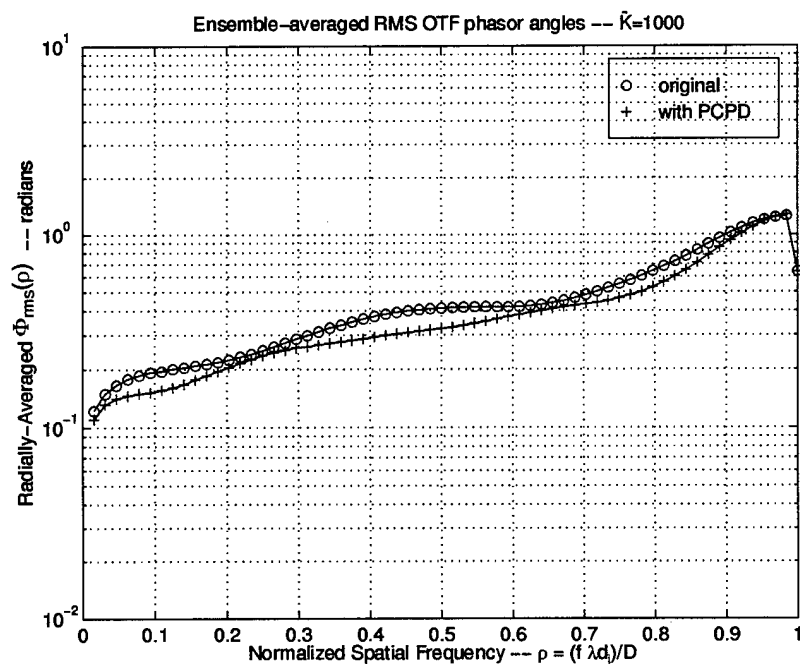


Figure 25. Case 1a — Zernike, low DOF. Radially-averaged profile of ensemble-averaged RMS OTF phasor angles before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$.

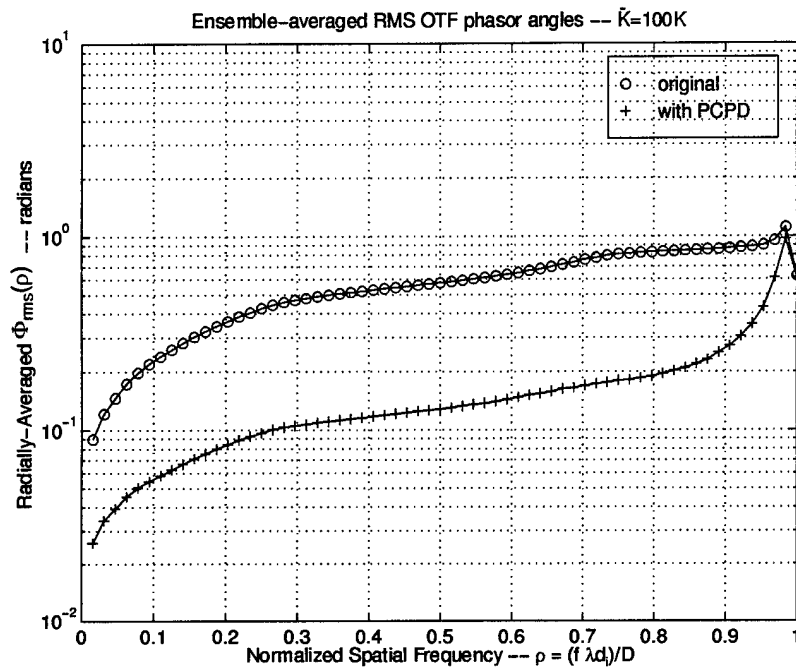


(a)

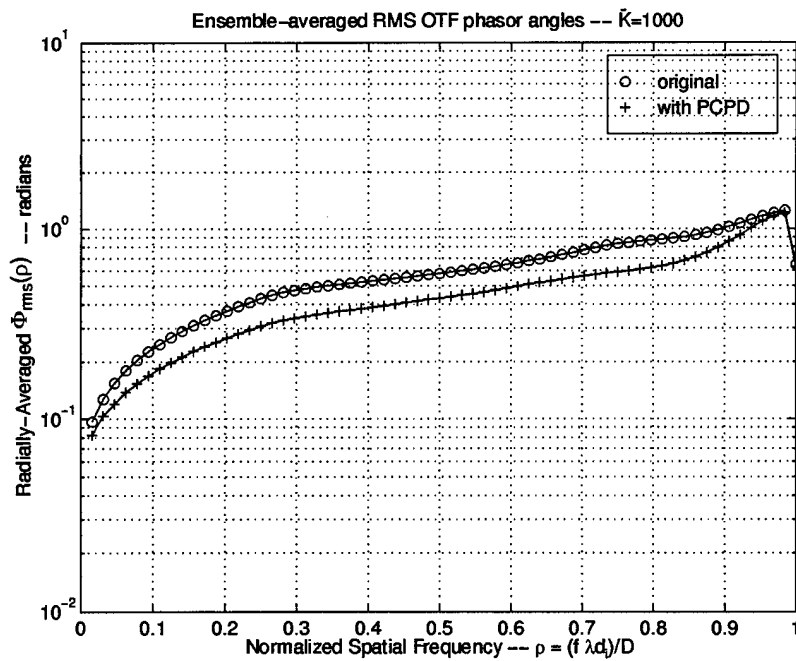


(b)

Figure 26. Case 2a — Zernike, high DOF. Radially-averaged profile of ensemble-averaged RMS OTF phasor angles before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$.

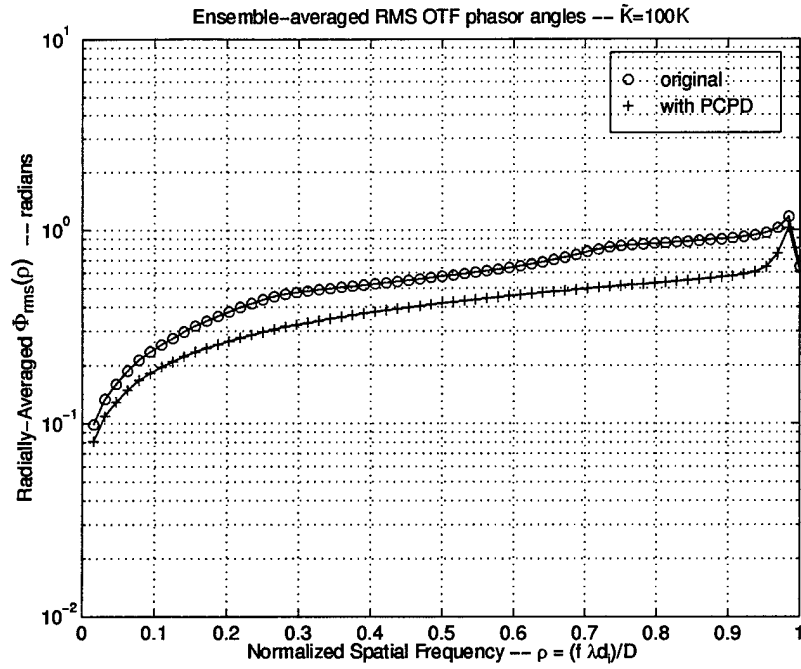


(a)

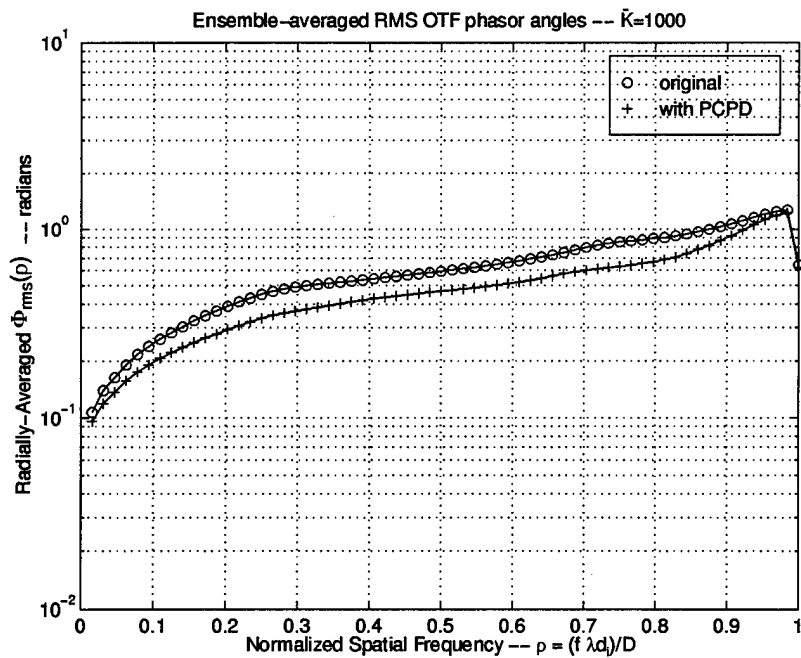


(b)

Figure 27. Case 1b — NGST, low DOF. Radially-averaged profile of ensemble-averaged RMS OTF phasor angles before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$.



(a)



(b)

Figure 28. Case 2b — NGST, high DOF. Radially-averaged profile of ensemble-averaged RMS OTF phasor angles before and after PCPD processing. (a): $\bar{K} = 10^5/\text{image}$; (b): $\bar{K} = 10^3/\text{image}$.

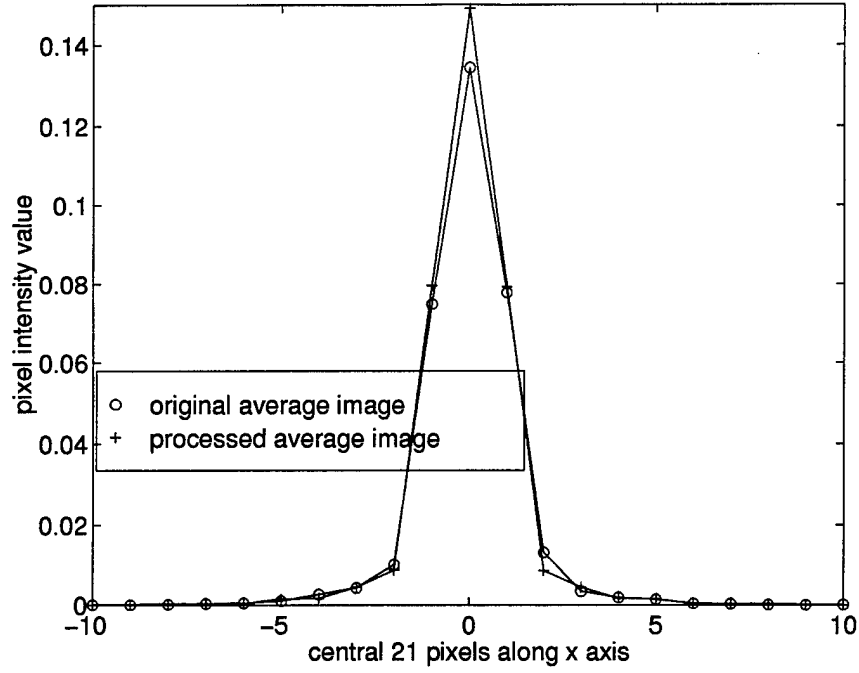


Figure 29. Average point source image slices before and after phase filtering using phase diversity WFS data, for the ensemble of 50 images, as described in the text. These data come from the $\bar{K} = 10^5$ NGST case 1, with average RMS aberration of $\lambda/10$ due to 8 petal piston random misalignment errors.

PCPD Strehl Improvement

Figure 29 demonstrates a case where this phase filtering has been accomplished based on the OTF estimate obtained via the phase diversity technique. The figure shows x -axis slices of the central 21 pixels of the average point-source images from NGST case 1. Recall that this is the low-DOF case with 100,000 photons per point source image. The original images are formed on a 64×64 array, with 2 pixels per λ/D . The central pixel intensity — essentially the unnormalized Strehl ratio — for the phase filtered ensemble average is 10% greater than that for the unprocessed ensemble average. The plot also shows a slightly tighter PSF, corresponding to more energy being concentrated in the central pixel.

4.5 Conclusion

This chapter presented results of a simulation study designed to investigate the performance of a least-squares phase diversity wavefront sensing technique. This effort was directed specifically towards investigation of the Gonsalves phase diversity wavefront sensing technique utilizing a straightforward regularization scheme. Simulation experiments exploring the photon-limited WFS performance of the technique were accomplished under a variety of wavefront sensing and aberration scenarios, specifically intended to approximate conditions of interest under the space-based (*i.e.* weakly-aberrated, nearly diffraction-limited) point source imaging scenario. Both a monolithic mirror with Zernike aberrations, and an NGST style mirror with zonal segment misalignment aberrations were simulated under 2 different conditions and 3 different light levels, giving a total of 12 different space telescope imaging scenarios.

Under the lowest-light (average of 1000 photons per image), highest degree-of-freedom case (26 segmented telescope misalignment parameters), the estimated wavefront diverged from the truth model answer in 20% of the ensemble realizations, yielding WFS errors that were incorrect by a few orders of magnitude. This result suggests that a certain minimum light-level threshold is required in order to estimate this many piston and tilt parameters. A result like this is not unexpected since the Gonsalves metric is the Gaussian log-likelihood function, while dim photon-limited point source images are not accurately modelled as Gaussian random processes.

For the other 11 scenarios, phase diversity WFS accuracy follows the expected downward trends as average light levels decrease and the number of aberration parameters to be estimated increases. The average RMS WFS estimation errors may be small in absolute terms, but when compared to the original RMS aberrations of

the nearly diffraction-limited pupils, the errors in phase diversity aberration estimation can be relatively large in comparison. This is most noticeably true in the cases of low light and/or high number of aberration parameters. Given an average aberration strength of 10% of a wave RMS, the worst case average WFS error was 5.7% of a wave RMS.

Using these WFS estimates in the phase deconvolution scheme of chapter 2 showed analogous improvements in the image quality metrics of Strehl ratio, and image spectral SNR and phase error.

The types of results presented in this chapter represent a significant addition to the body of experience with this relatively new WFS technique. Whether the WFS estimates are to be incorporated into an adaptive optical mirror phasing scheme, or are used in post-processing, it appears that PDWFS estimates can provide significant improvement in imaging performance for weakly-aberrated imaging systems, such as space telescopes, both segmented and monolithic.

V. Cramér-Rao analysis of phase diversity imaging

5.1 Introduction

The analysis presented in this chapter combines and expands upon expressions found in references (18, 62), in order to give the Cramér-Rao lower bound (CRLB) on the phase diversity estimation of aberration polynomial coefficients. This analysis specifically applies to a set of photon-limited images from which the aberration information is to be extracted using some unspecified estimator. The analysis allows for the isoplanatic imaging of an arbitrary extended object irradiance distribution.

Whenever a parameter is being estimated using random data, the accuracy of the estimate is fundamentally limited by the randomness of the data. This straightforward, intuitive idea is expressed more quantitatively in the Cramér-Rao theorem. The smallest possible inaccuracy or variance for a given estimation problem is given by the CRLB. The estimator variance lower bound for the phase diversity problem is derived here under the assumption that the aberration is fixed but unknown, where the only randomness in the problem is described by photon statistics.

Section 5.2 gives the basic mathematical background, and section 5.3 derives the analytical form for the phase diversity CRLB. It turns out that the CRLB for estimating aberration coefficients is dependent on the actual, underlying values of the aberration coefficients being estimated, so in section 5.4 we give some phase diversity CRLB values for a number of representative aberration cases by way of computer simulation.

5.2 Notation and preliminary expressions

This chapter begins by developing some preliminary expressions needed to derive the phase diversity CRLB. Specifically, the analysis will require partial derivatives of image domain quantities with respect to pupil-domain quantities. These partial derivative expressions will be used in developing the CRLB in section 5.3.

5.2.1 Fourier optics revisited. The reader will recall that the linear, isoplanatic, Fourier optical imaging model has already been presented in chapter 2, based on refs. (29, 22, 72). Some of these results are presented again here, in more detail and slightly different form. The summary of Fourier optics presented again below will facilitate the development of the partial derivative expressions, which will be needed in order to proceed with the CRLB derivation.

Recall from equation 6 of chapter 2 that the aperture of the hypothetical imaging system is denoted by the binary indicator function, given this time, with the symbol A , as

$$(71) \quad A(\vec{u}) = \begin{cases} 1 & \text{for } \vec{u} \in \text{pupil} \\ 0 & \text{otherwise.} \end{cases}$$

Notice the distinction between pupil coordinates \vec{u} , and later, image plane coordinates, \vec{x} . In order to finally arrive at a point-spread function (PSF) of unit integrated volume below, we pre-normalize the aperture, and use a weighted aperture function $W(\vec{u})$ that is a scaled version of $A(\vec{u})$, such that

$$(72) \quad \int_{\vec{u}} W^2(\vec{u}) = 1.$$

This means that if the total area of the aperture is denoted by a , then the weighted aperture is defined as

$$(73) \quad W(\vec{u}) = \begin{cases} 1/\sqrt{a} & \text{for } \vec{u} \in \text{pupil} \\ 0 & \text{otherwise.} \end{cases}$$

More detail on this pupil normalization is found in the appendix.

The generalized pupil function, which is a scaled version of equation 7 of chapter 2, is again defined using the phase angles of complex exponentials in the pupil to represent optical path differences (OPDs) caused by aberrations $\phi(\vec{u}; \vec{\alpha})$:

$$(74) \quad g(\vec{u}; \vec{\alpha}) = W(\vec{u}) \exp[j\phi(\vec{u}; \vec{\alpha})].$$

Note that the vector $\vec{\alpha}$ again refers to the the phase screen coefficients in some basis set, such as the Zernike polynomials, Z_j . The Fourier transform of the generalized pupil function is the coherent PSF:

$$(75) \quad G(\vec{x}; \vec{\alpha}) = \text{FT} \left[g(\vec{u}; \vec{\alpha}) \right].$$

The incoherent PSF is then given by

$$(76) \quad h(\vec{x}; \vec{\alpha}) = |G(\vec{x}; \vec{\alpha})|^2 = G(\vec{x}; \vec{\alpha}) G^*(\vec{x}; \vec{\alpha}),$$

where this quantity is automatically normalized to unit integrated volume via equation 73. The deterministic, isoplanatic, Fourier optics image is a convolution of the PSF with the object irradiance distribution $o(\vec{x})$, or

$$(77) \quad \lambda(\vec{x}; \vec{\alpha}) = o(\vec{x}) * h(\vec{x}; \vec{\alpha})$$

where, to be consistent with statistical optics notation conventions (29), the symbol i is replaced with λ to denote the noiseless image. Since this effort is concerned solely with detection in the form of photon-counting, the integral of the received image is equated to the sum total number of photons counted, designated as \overline{K} :

$$(78) \quad \int_{\vec{x}} \lambda(\vec{x}; \vec{\alpha}) d\vec{x} = \overline{K}.$$

Let the quantity $o_{norm}(\vec{x})$ be defined as a scaled version of $o(\vec{x})$ such that

$$(79) \quad \int_{\vec{x}} o_{norm}(\vec{x}) d\vec{x} = 1.$$

The image quantity can then be written as

$$(80) \quad \lambda(\vec{x}; \vec{\alpha}) = \overline{K} o_{norm}(\vec{x}) * h(\vec{x}; \vec{\alpha})$$

$$(81) \quad = \overline{K} o_{norm}(\vec{x}) * \left(G(\vec{x}; \vec{\alpha}) G^*(\vec{x}; \vec{\alpha}) \right).$$

This normalization explicitly presents the functional dependence of subsequent Cramér-Rao quantities on the total image photon level, \overline{K} , a dependence which is obscured by the development found in the corresponding references (18, 62). In the semi-classical model of photodetection (29) which we invoke in section 5.3, note that the quantity \overline{K} is actually the expected value of image photocount, K , a random quantity.

5.2.2 Preliminary expressions: partial derivatives. The CRLB analysis which follows will require expressions for the derivatives of the image quantity λ with respect to the individual pupil aberration coefficients. This section follows the derivation of such partial derivatives found in reference (18), generalizing to account for isoplanatic imaging of an extended source object—*i.e.* the convolution operation

of equation 77. Proceeding in stages starting with equation 80, the partial derivative can be written as:

$$(82) \quad \frac{\partial \lambda(\vec{x}; \vec{\alpha})}{\partial \alpha_k} = \overline{K} o_{norm}(\vec{x}) * \frac{\partial h(\vec{x}; \vec{\alpha})}{\partial \alpha_k},$$

where the linearity of the convolution and differentiation operations allows the order of the operations to be exchanged. Applying the product rule to equation 76 gives:

$$(83) \quad \begin{aligned} \frac{\partial h(\vec{x}; \vec{\alpha})}{\partial \alpha_k} &= \frac{\partial G^*(\vec{x}; \vec{\alpha})}{\partial \alpha_k} G(\vec{x}; \vec{\alpha}) + \frac{\partial G(\vec{x}; \vec{\alpha})}{\partial \alpha_k} G^*(\vec{x}; \vec{\alpha}) \\ &= 2 \operatorname{Real} \left[\frac{\partial G(\vec{x}; \vec{\alpha})}{\partial \alpha_k} G^*(\vec{x}; \vec{\alpha}) \right]. \end{aligned}$$

This equation in turn depends on the derivative of the coherent PSF:

$$(84) \quad \begin{aligned} \frac{\partial G(\vec{x}; \vec{\alpha})}{\partial \alpha_k} &= \frac{\partial}{\partial \alpha_k} \left(\operatorname{FT} \left[g(\vec{u}; \vec{\alpha}) \right] \right) \\ &= \frac{\partial}{\partial \alpha_k} \left(\operatorname{FT} \left[W(\vec{u}) \exp \left\{ j \left(\sum_{i=1}^J \alpha_i Z_i(\vec{u}) \right) \right\} \right] \right) \\ &= \frac{\partial}{\partial \alpha_k} \left(\int_{\vec{u}} \left[W(\vec{u}) \exp \left\{ j \left(\sum_{i=1}^J \alpha_i Z_i(\vec{u}) \right) \right\} \exp \{ j 2 \pi \vec{u} \cdot \vec{x} \} \right] d\vec{u} \right). \end{aligned}$$

The full expansion shown in equation 84 enables us to isolate the single term of the integrand which contains α_k , and the derivative of the coherent PSF is given as

$$(85) \quad \begin{aligned} \frac{\partial G(\vec{x}; \vec{\alpha})}{\partial \alpha_k} &= \int_{\vec{u}} \left(g(\vec{u}; \vec{\alpha}) \left(j Z_k(\vec{u}) \right) \exp \{ j 2 \pi \vec{u} \cdot \vec{x} \} \right) d\vec{u} \\ &= j \operatorname{FT} \left[g(\vec{u}; \vec{\alpha}) Z_k(\vec{u}) \right]. \end{aligned}$$

Applying this result in equation 83 gives

$$\begin{aligned}
 \frac{\partial h(\vec{x}; \vec{\alpha})}{\partial \alpha_k} &= 2 \operatorname{Real} \left[\frac{\partial G(\vec{x}; \vec{\alpha})}{\partial \alpha_k} G^*(\vec{x}; \vec{\alpha}) \right] \\
 &= 2 \operatorname{Real} \left[j \left(\operatorname{FT} \left[g(\vec{u}; \vec{\alpha}) Z_k(\vec{u}) \right] \right) G^*(\vec{x}; \vec{\alpha}) \right] \\
 (86) \qquad &= -2 \operatorname{Imag} \left[\left(\operatorname{FT} \left[g(\vec{u}; \vec{\alpha}) Z_k(\vec{u}) \right] \right) G^*(\vec{x}; \vec{\alpha}) \right].
 \end{aligned}$$

The desired partial derivative is then given as

$$(87) \quad \frac{\partial \lambda(\vec{x}; \vec{\alpha})}{\partial \alpha_k} = -2 \overline{K} o_{norm}(\vec{x}) * \operatorname{Imag} \left[\left(\operatorname{FT} \left[g(\vec{u}; \vec{\alpha}) Z_k(\vec{u}) \right] \right) G^*(\vec{x}; \vec{\alpha}) \right].$$

This result will be used in the Cramér-Rao lower bound analysis shown next.

5.3 Phase diversity Cramér-Rao lower bound derivation

We now use the expressions developed above, along with statistical optics concepts, to derive the minimum possible estimator variance for the phase diversity wavefront sensing technique, slightly generalizing the analysis shown in reference (18) to account for the phase diversity imaging situation.

5.3.1 Cramér-Rao concept. The Cramér-Rao inequality is a fundamental concept of estimation theory, written symbolically (43) as

$$(88) \qquad \operatorname{var}(\tilde{\alpha}_k) \geq [\mathbf{F}^{-1}(\vec{\alpha})]_{kk},$$

where the tilde denotes an unbiased estimate of α_k , and $[\mathbf{F}(\vec{\alpha})]$ represents the Fisher information matrix of the parameter vector $\vec{\alpha}$. In words, the inequality states that the mean-squared error (MSE) of an unbiased estimate of some element of $\vec{\alpha}$ is

bounded from below by the corresponding diagonal element of a Fisher information matrix inverse. The Fisher matrix is defined as the negative of the expected value of the second partial derivative matrix the log-likelihood function of the measurement data. This second partial derivative is taken with respect to all of the pairs the estimated parameters. This is shown symbolically as

$$(89) \quad \mathbf{F}(\vec{\alpha}) = -\mathbf{E} \left[\mathbf{F}_o(\vec{\alpha}) \right],$$

and the observed Fisher information matrix is given by

$$(90) \quad [\mathbf{F}_o(\vec{\alpha})]_{jk} = \frac{\partial^2 L(d; \vec{\alpha})}{\partial \alpha_j \partial \alpha_k}$$

where $L(d; \vec{\alpha})$ is the log-likelihood function of the data, described below. The Fisher information matrix is therefore a *Hessian* matrix, which means that the i, j th element of the matrix is the second partial with respect to i, j th aberration parameters.

5.3.2 Phase diversity semi-classical photodetection model. The noisy phase diversity data are symbolized by $\{d_n(\vec{x})\}$, and are noisy realizations of $\{\lambda_n(\vec{x})\}$, by virtue of the fact that the d_n 's contain a finite, random number of photoevents, K . Notice also that, depending on the beamsplitting arrangement between the various phase diversity images, each of the N diversity images may receive different percentages of the incoming total irradiance. Since the phase estimates are based on noisy image data, the accuracy of the estimate is fundamentally limited by the randomness of the measured images, and it is the probability density of this data set that is of interest here.

The symbol $d_n(\vec{x})$ represents the random variable for the number of photons striking a detector at location \vec{x} in the n th diversity image. The probability density

function of this variable, $\text{PDF}[d_n(\vec{x})]$, follows the Poisson distribution,

$$(91) \quad \text{PDF} \left[d_n(\vec{x}) | o | \vec{\alpha} \right] = \frac{(\lambda_n(\vec{x}; \vec{\alpha}))^{d_n(\vec{x})} \exp \{ -\lambda_n(\vec{x}; \vec{\alpha}) \}}{d_n(\vec{x})!}.$$

The quantity $\lambda_n(\vec{x}; \vec{\alpha})$ is often referred to as the mean, or “infinite-signal” image, or the rate function of the random photocount PDF, because of the well-known property of Poisson processes,

$$(92) \quad E[d_n(\vec{x})] = \lambda_n(\vec{x}; \vec{\alpha}),$$

where $E[\dots]$ symbolizes the expectation operator.

Another statistical property of the semi-classical model is that photoevents at different locations \vec{x} , and in different diversity images, are statistically independent. Statistical independence allows the density of the entire noisy diversity image data set to be written simply as a product of the individual event probabilities, or

$$(93) \quad \text{PDF} \left[\{d_n(\vec{x})\} \right] = \prod_{n=1}^N \prod_{\vec{x}} \text{PDF} \left[d_n(\vec{x}) \right],$$

where the set brackets $\{\dots\}$ refer to a particular set of phase diverse image data collected across all diversity images d_n and across all image locations \vec{x} . The logarithm of the above PDF gives the desired log-likelihood function:

$$(94) \quad \begin{aligned} L(d; \vec{\alpha}) &= \ln \left(\prod_{n=1}^N \prod_{\vec{x}} \text{PDF} \left[d_n(\vec{x}) \right] \right) \\ &= \sum_{n=1}^N \sum_{\vec{x}} \ln \left(\text{PDF} \left[d_n(\vec{x}) \right] \right) \\ &= \sum_{n=1}^N \sum_{\vec{x}} \left(-\lambda_n(\vec{x}; \vec{\alpha}) + d_n(\vec{x}) \ln(\lambda_n(\vec{x}; \vec{\alpha})) - \ln(d_n(\vec{x})!) \right) \end{aligned}$$

Recall that the photocount random variable $d_n(\vec{x})$ is not dependent on $\vec{\alpha}$, and further, that the partial derivatives of $L(d; \vec{\alpha})$ with respect to $\vec{\alpha}$ are the actual quantities of interest. Since the partial derivative of $\ln(d_n(\vec{x})!)$ with respect to $\vec{\alpha}$ will be zero, this last term in the summand of equation 94 is discarded, leaving

$$(95) \quad L(d; \vec{\alpha}) = \sum_{n=1}^N \sum_{\vec{x}} \left(-\lambda_n(\vec{x}; \vec{\alpha}) + d_n(\vec{x}) \ln(\lambda_n(\vec{x}; \vec{\alpha})) \right),$$

from which the Fisher information matrix will be derived.

5.3.3 Phase diversity Fisher information matrix. The partial derivatives needed for equation 90 can now be found, beginning with the first partial derivative of equation 95:

$$(96) \quad \frac{\partial L(d; \vec{\alpha})}{\partial \alpha_k} = \sum_{n=1}^N \sum_{\vec{x}} \frac{\partial \lambda_n(\vec{x}; \vec{\alpha})}{\partial \alpha_k} \left[\frac{d_n(\vec{x})}{\lambda_n(\vec{x}; \vec{\alpha})} - 1 \right].$$

This expression is an extension of the corresponding expression found in reference (18)—the image under consideration has been generalized from a point source to an extended image, and an extra summation is added to account for the multiple diversity images. These extensions will allow some interesting aspects of the phase diversity problem to be explored, as discussed in section 5.4. Similar expressions appear in reference (62), related to the derivatives of photon-limited image likelihood functions with respect to pupil aberration parameters.

At this point, equation 96 can be used to verify that the PDF satisfies the regularity condition for the Cramér-Rao theorem, given in reference (43). This condition dictates that in order for the CRLB theorem to hold the following condition

must be satisfied:

$$(97) \quad E \left[\frac{\partial L(d; \vec{\alpha})}{\partial \vec{\alpha}} \right] = \vec{0},$$

the zero vector. Recalling equation 92, the quotient in equation 96 becomes

$$(98) \quad \begin{aligned} E \left[\frac{d_n(\vec{x})}{\lambda_n(\vec{x}; \vec{\alpha})} \right] &= \frac{1}{\lambda_n(\vec{x}; \vec{\alpha})} E \left[d_n(\vec{x}) \right] \\ &= \frac{\lambda_n(\vec{x}; \vec{\alpha})}{\lambda_n(\vec{x}; \vec{\alpha})} \end{aligned}$$

and so the expectation of equation 96 vanishes, and the regularity condition holds.

To find the second partial derivatives needed for equation 90, equation 96 is differentiated with respect to a second aberration coefficient α_k , written as

$$(99) \quad \begin{aligned} [\mathbf{F}_o(\vec{\alpha})]_{jk} &= \frac{\partial^2 L(d; \vec{\alpha})}{\partial \alpha_j \partial \alpha_k} \\ &= \sum_{n=1}^N \sum_{\vec{x}} \left(\left[\frac{\partial^2 \lambda_n(\vec{x}; \vec{\alpha})}{\partial \alpha_j \partial \alpha_k} \right] \left[\frac{d_n(\vec{x})}{\lambda_n(\vec{x}; \vec{\alpha})} - 1 \right] - \right. \\ &\quad \left. \left[\frac{\partial \lambda_n(\vec{x}; \vec{\alpha})}{\partial \alpha_j} \frac{\partial \lambda_n(\vec{x}; \vec{\alpha})}{\partial \alpha_k} \frac{d_n(\vec{x})}{(\lambda_n(\vec{x}; \vec{\alpha}))^2} \right] \right). \end{aligned}$$

The negative of the expected value of the above equation is actually the desired quantity. Again taking advantage of the property of Poisson shot processes shown in equation 92, it is easily seen that the expectation operator can be used to cause d quantities to become λ quantities, thus annihilating the first summand term of equation 99 vanish. By also slightly simplifying the second term of the summand

the following can be written:

$$\begin{aligned}
 [\mathbf{F}(\vec{\alpha})]_{jk} &= -\mathbf{E} \left[[\mathbf{F}_o(\vec{\alpha})]_{jk} \right] \\
 (100) \qquad &= \sum_{n=1}^N \sum_{\vec{x}} \left(\frac{\partial \lambda_n(\vec{x}; \vec{\alpha})}{\partial \alpha_j} \frac{\partial \lambda_n(\vec{x}; \vec{\alpha})}{\partial \alpha_k} \frac{1}{\lambda_n(\vec{x}; \vec{\alpha})} \right).
 \end{aligned}$$

Since no specific Fourier-optical aspects are yet embodied in equation 100, we note that it is essentially a re-statement of equation 2.62 of reference (76). This is a general result which is applicable to any estimation of a noisy parameter $\vec{\alpha}$ using noisy Poisson counting data.

Using the partial derivative results of the previous section, equation 87, the Fisher matrix is written as

$$\begin{aligned}
 [\mathbf{F}(\vec{\alpha})]_{jk} &= \sum_{n=1}^N \sum_{\vec{x}} \frac{4(\overline{K_n})^2}{\lambda_n(\vec{x}; \vec{\alpha})} o_{norm}(\vec{x}) * \left(\text{Imag} \left\{ G_n^*(\vec{x}; \vec{\alpha}) \text{FT} \left[g_n(\vec{u}; \vec{\alpha}) Z_j(\vec{u}) \right] \right\} \right) \times \\
 &\quad o_{norm}(\vec{x}) * \left(\text{Imag} \left\{ G_n^*(\vec{x}; \vec{\alpha}) \text{FT} \left[g_n(\vec{u}; \vec{\alpha}) Z_k(\vec{u}) \right] \right\} \right) \\
 &= \sum_{n, \vec{x}} \frac{4(\overline{K_n})^2}{\overline{K_n} o_{norm}(\vec{x}) * h_n(\vec{x}; \vec{\alpha})} \times \\
 &\quad \prod_{a=j,k} o_{norm}(\vec{x}) * \left(\text{Imag} \left\{ G_n^*(\vec{x}; \vec{\alpha}) \text{FT} \left[g_n(\vec{u}; \vec{\alpha}) Z_a(\vec{u}) \right] \right\} \right) \\
 &= \sum_{n, \vec{x}} \overline{K_n} \frac{4}{o_{norm}(\vec{x}) * h_n(\vec{x}; \vec{\alpha})} \times \\
 (101) \qquad &\quad \prod_{a=j,k} o_{norm}(\vec{x}) * \left(\text{Imag} \left\{ G_n^*(\vec{x}; \vec{\alpha}) \text{FT} \left[g_n(\vec{u}; \vec{\alpha}) Z_a(\vec{u}) \right] \right\} \right)
 \end{aligned}$$

Equation 101 is the final desired result, the Fisher information matrix for the extended-object, multiple-image, phase diverse data set.

Note that the basis set $\{Z_j(\vec{u})\}$ has been taken to mean the Zernike polynomial basis set (58) for circular pupils, a basis set which will be used in the numerical

demonstrations in the latter section of this chapter. However, the selection of basis set was not crucial to the derivation. For the theoretical development given above, the shape of the aperture and the actual aberration basis set could have remained unspecified. Alternatively, the set symbolized by $\{Z_j(\vec{u})\}$ can be interpreted as denoting any given pupil basis set, up to and including discrete pupil domain delta functions. Before proceeding, it is also important to point out that the CRLB is only a theoretical lower bound for the *unbiased* estimation problem, and it is at least theoretically possible that there exists some biased estimator that exhibits better estimation performance (18, 43).

The Cramér-Rao lower bound on the mean-squared error of an aberration parameter estimate can now, in principle, be calculated. First, the Fisher information matrix is populated via equation 101. This matrix is then inverted, and the desired lower bounds are found in the appropriate diagonal elements of this inverse. One practical problem with this scheme is that some of the items needed to evaluate equation 101 depend directly upon the aberrations $\vec{\alpha}$ that are being estimated. Strictly speaking, there would be no reason to find the lower bound for an estimate if the quantity being estimated is already known exactly, as assumed in equation 101. Similarly, the original object irradiance distribution, $o_{norm}(\vec{x})$, is also not known exactly in most astronomical imaging situations. Once again, there would most likely be no reason to carry out the observation if the object were already known.

Note that the object and aberration could be specified *a priori* in contrived situations such as controlled laboratory experiments or computer imaging simulations. A CRLB analysis utilizing this latter approach might involve finding the Cramér-Rao lower bounds for a large number of objects and aberrations. For instance, an ensemble of phase screens belonging to certain statistical class of random aberrations

could be generated and used in equation 101, along with a certain type of imaging target object distribution. Estimation-theoretic conclusions might then be drawn from such an analysis which are approximately valid for aberrations of that certain statistical class. This approach to Cramér-Rao analysis is investigated in the next section of this chapter.

5.4 *Simulation examples*

To proceed with numerical evaluation of the expressions of the previous section, the computer imaging simulation from previous chapters was adapted to the numerical evaluation of equation 101 for numerically generated pupil phase screens and the corresponding PSFs and images. For this section, the pupil phases are again written to a 128×128 pixel grid with a circular pupil aperture 64 pixels in diameter located at the center of the array. The quantities needed for equation 101 are numerically obtained using Fourier optical relations and Fast Fourier transforms (FFTs). The pupil phases are again specified in terms of the Zernike polynomials as ordered and normalized by Noll (58).

Average Lower Bounds—Monte-Carlo analysis of CRLBs

As seen in the previous section, particularly equation 101, the Cramér-Rao lower bound for an aberration estimation problem depends upon the *actual*, underlying values of the aberration being estimated, and the bounds generally change as the aberrations change. Thus it is difficult to make any meaningful general quantitative statements about the behavior of the CRLB without reference to a given, specific aberration example. To overcome this difficulty, a "Monte-Carlo" style CRLB analysis is presented. This analysis proceeds by averaging Cramér-Rao lower bounds across an ensemble of random pupils of a certain statistical class, as described below.

The purpose of this averaging is to obtain a more generic, CRLB-related quantity, which is more broadly applicable, at least illustratively.

The intuitive idea of interpreting a single CRLB value as a figure of merit has been seen previously in the literature, refs. (18, 43), where lower the CRLBs are presumed to be consistent with better estimation techniques. The idea of an “ensemble average” Cramér-Rao lower bound may at first glance seem controversial since, obviously, an “average lower bound” is no longer a true lower bound. However, the Cramér-Rao lower bound can also be re-interpreted as the mean-squared error (MSE) of some theoretical minimum-variance unbiased estimator (MVUE), an estimator which may or may not exist (43). The idea of ensemble averages of pupil averaged RMS phase errors—including phase estimation errors—is a common feature of much of the literature of adaptive optics for example, such as references (19, 58, 78, 80), and was also used to summarize the results of chapter 4. The approach presented below is therefore interpreted as the MSE performance of a hypothetical WFS estimator. Instead of averaging lower bounds, which may sound nonsensical, the mean squared error (MSE) of the hypothetical MVUE is being estimated by way of ensemble averaging.

The concept of the Monte-Carlo analysis used here is depicted in block-diagram form in figure 30. The MVUE MSE for each randomly simulated phase screen is first determined via the trace of the inverse of the corresponding Fisher information matrix. Note that the MVUE MSE of the coefficients corresponds to the MVUE MSE of the actual pupil phases for the case of Noll-modified Zernike polynomials, because of basis set orthonormality (58). These N minimum mean-squared phase errors quantities are averaged across the ensemble to obtain the average CRLB metric of interest. If the individual CRLB of the i th pupil is referred to as ϵ_i^2 , then the

Monte-Carlo average metric discussed here is represented by

$$(102) \quad \text{Data Point} = \frac{\sum_{i=1}^N \epsilon_i^2}{N} = \langle \epsilon^2 \rangle$$

5.4.1 Optimal defocus example. For each of the data shown in figure 31, fifty pupil phase screens were generated, with an ensemble-averaged, pupil-averaged RMS aberration phase of one-quarter of a wave. This was accomplished through random number generation of Zernike coefficients 4 through 11. All 8 of these coefficients were generated as independent, identically distributed Gaussian random vectors. The requisite quantities for equation 101 are each determined Fourier optically, and the Hessian Fisher matrix is populated by numerical quadrature. For this simulated pupil case the Fisher matrix is an 8×8 matrix corresponding to Zernike coefficients 4 through 11.

The same set of 50 underlying pupil phase aberrations was used for all of the simulated cases summarized by figure 31. The simulated total \overline{K} quantity was normalized to unity with 50 percent of the available light to each of two simulated images. The phase diversity imaging configuration of figure 1 is assumed in the calculation of equation 101. The amount of defocus was allowed to vary in the second image by adding the desired amount of Zernike mode 4 to the diversity image pupil phase. This diversity defocus parameter is the quantity denoted on the x-axis of the plot. Point-source imaging was simulated for these cases. For all of the CRLB data given in the rest of this chapter, the MSEs are given normalized for 1 total photon. For an actual total photocount of \overline{K} , over all diversity images, the Cramér-Rao answers given here must be divided by \overline{K} to obtain the actual MSE, which is in units of radians squared.

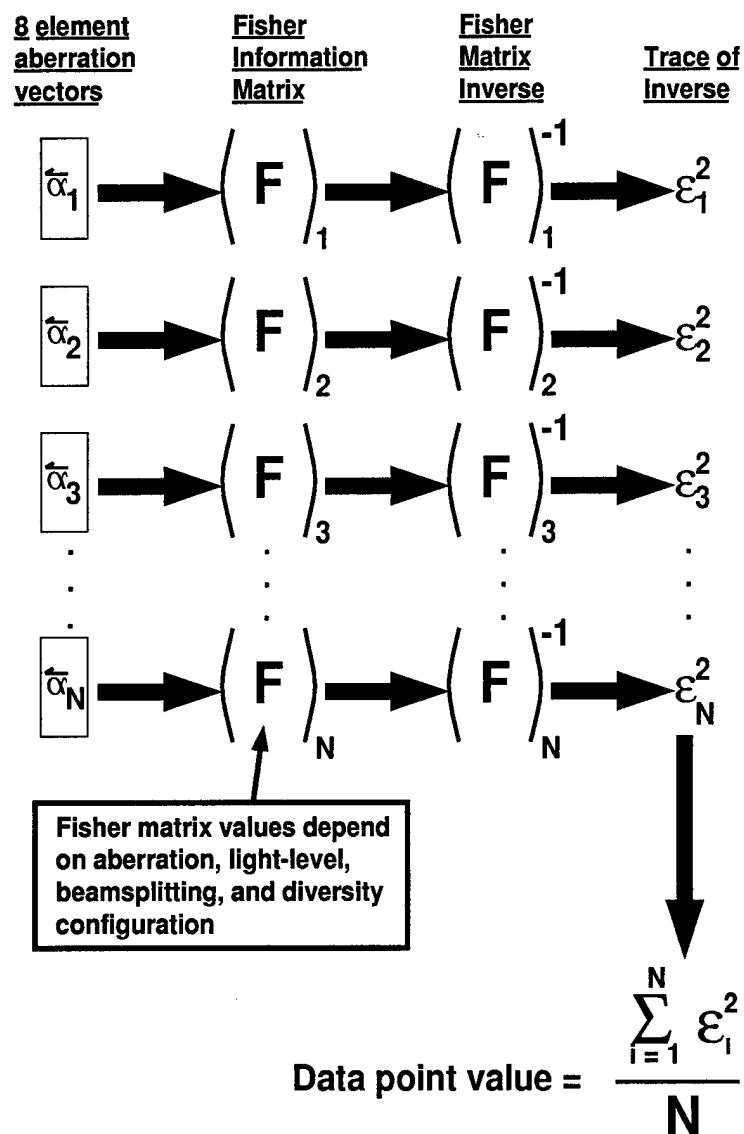


Figure 30. Graphical explanation of the procedure used to obtain the Cramér-Rao figure-of-merit metric values of this section. This general procedure is justified and explained in the text.

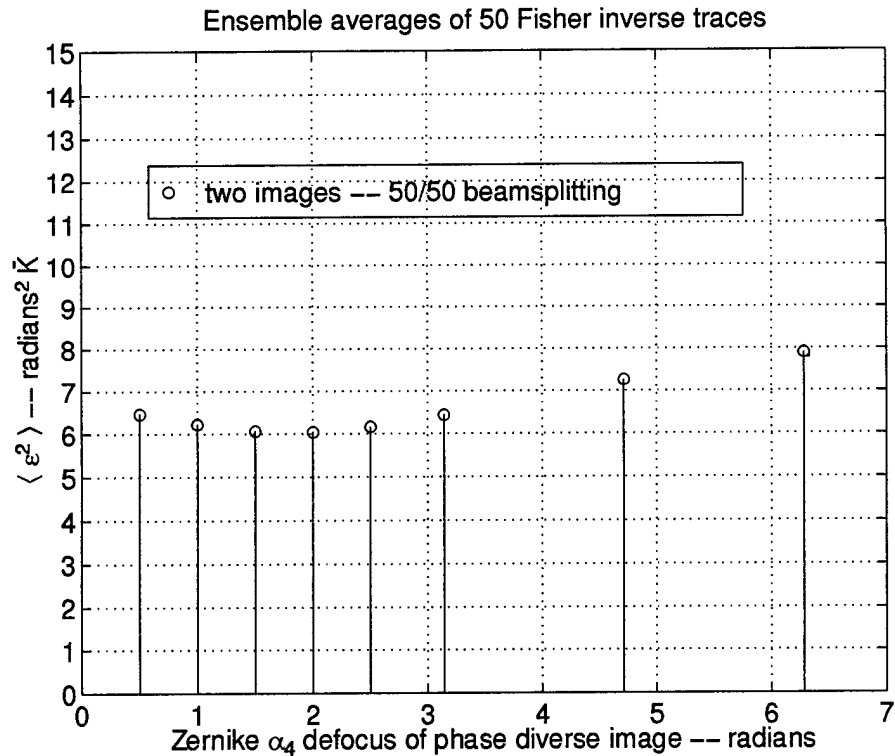


Figure 31. Ensemble average of pupil-averaged theoretical minimum mean-squared phase diversity estimation error for 50 quarter-wave RMS aberrated pupils. Data points represent the ensemble averages of the traces of the inverses of the 50 Fisher information matrices. Fisher matrices calculated with 50/50 beamsplitter, total photocount normalized to $\bar{K} = 1$. Traditional phase diversity setup with one image in focus and one image out-of focus by the amount indicated on the x-axis. Point-source imaging is modelled.

Any given data point of figure 31 is of interest by itself, since it gives an indication of the average minimum squared error of the phase diversity MVUE, normalized by photocount. But the plot also allows comparison of the various MVUE error values as the defocus of the second diversity image is allowed to vary. By these criterion, the trend appears to indicate that there may be an optimum diversity defocus distance, corresponding to approximately 2 radians of Zernike mode 4.

5.4.2 Curvature sensing example. Curvature sensing (67) is another wavefront sensing technique that is amenable to the type of Cramér-Rao analysis shown above. As shown schematically in figure 32, the curvature sensing WFS method involves the collection of two symmetrically defocused images. Roddier (67) has shown that the difference between the two symmetrically defocused images contains information regarding the pupil phase screen's two-dimensional Laplacian. Solution of Laplace's equation is the post-detection technique Roddier uses to extract a wavefront estimate. But equation 35 implies that these two symmetrically defocused images could very well also serve as a phase diversity data set and could be processed as any other PDWFS image set, using equation 47 of chapter 3, instead of Roddier's method.

Note that Cramér-Rao analysis provides a fundamental theoretical lower bound on error, regardless of the specific technique used to process the photon-limited data, be it inverting a Laplacian operator or Gonsalves function minimization. Thus the analysis of the previous subsection can be repeated for the case of a symmetrical defocused pair of images, regardless of the ultimate estimation paradigm to be used. Incidentally, Roddier's method is limited to point-source images, while PDWFS is not.

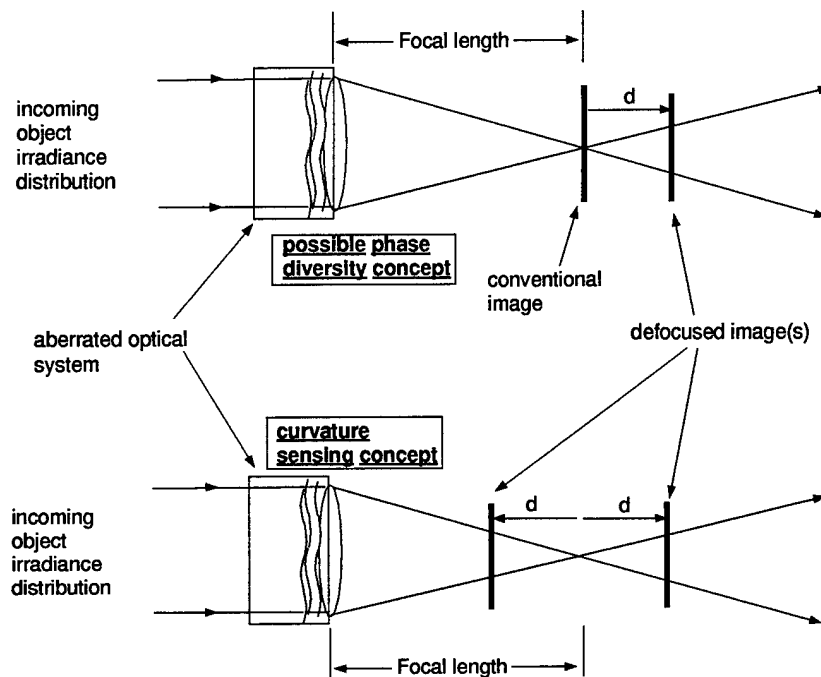


Figure 32. Schematic comparison of the curvature sensing and phase diversity techniques, in terms of the types of collected images.

Moreover, it may be of interest to attempt to use symmetrically defocused image pairs in combination with the phase diversity style of post-processing (15). As stated previously, there is no compelling fundamental reason why such image pairs cannot be used in this manner. The question then becomes one of whether this image collection scheme has any information-theoretic advantage over the traditional PDWFS configuration of figure 1. The results presented in figure 33 are an example of a type of analysis that could be useful in trying to answer such a question, when compared to corresponding results like those shown in figure 31.

The same ensemble of 50 quarter-wave aberrating phase screens used to create figure 31 were again used in making figure 33. In these curvature sensing cases, however, a pair of images was created by perturbing the phase screen by both $\pm\alpha_4$ radians of defocus. This is the x -axis variable of figure 33. For each pair of diversity

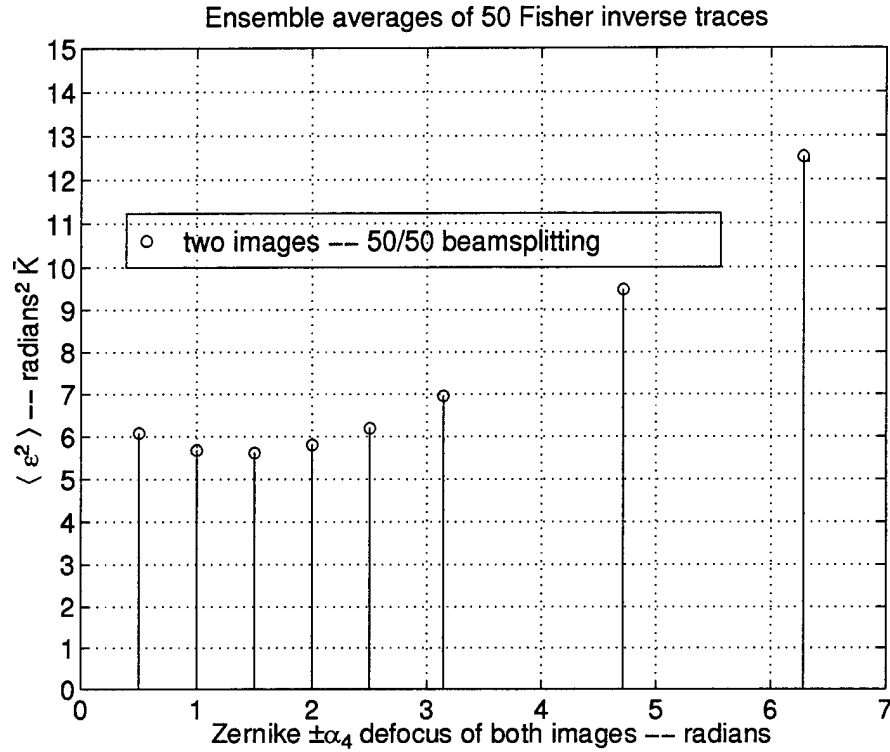


Figure 33. Ensemble average of pupil-averaged theoretical minimum mean-squared curvature sensing estimation error for 50 quarter-wave RMS aberrated pupils. Data points represent the ensemble averages of the traces of the inverses of the 50 Fisher information matrices. Fisher matrices calculated with 50/50 beamsplitter, total photocount normalized to $\bar{K} = 1$. Traditional phase diversity setup with one image in focus and one image out-of focus by the amount indicated on the x-axis. Point-source imaging is modelled.

Table 2. The summary of results of the various phase diversity CRLB experiments. Photocounts normalized to unity.

experiment no.	description	$\langle \epsilon^2 \rangle$
1	single pt. src. image	6.54 rad ²
2	beamsplit 30/70	6.37 rad ²
3	beamsplit 40/60	6.17 rad ²
4	3 phase div. images	6.48 rad ²
5	satellite object	227.5 rad ²
6	cone filtering	17.1 rad ²
7	chapter 4 case 1a	11.18 rad ²
8 best fig 31	section 5.4.1	6.05 rad ²
9 best fig 33	section 5.4.2	5.62 rad ²

images, the calculation of the Fisher information matrix again proceeded as before, via equation 101, with \bar{K} again normalized to unity. The ensemble-averaged MSEs were found for each curvature sensing configuration using the procedure of figure 30.

Figure 33 shows that out of the various configurations tested, a symmetric defocus of $+/- 1.5$ radians appears to be the best curvature sensing configuration in an MVUE MSE sense, for this group of aberrating phase screens. Compare this to the corresponding best phase diversity configuration with 2 radians of diversity, as determined from figure 31. The best curvature sensing configuration appears to exhibit 7 percent less MSE than any standard phase diversity configuration.

5.4.3 Other experiments. The results of various other CRLB experiments are summarized in table 2. These experiments are discussed in greater detail below. All except experiment 7 utilize the same ensemble of 50 underlying, quarter wave RMS aberrating phase screens for the averaging. The last two items give the best results shown in figures 31 and 33. The same type of analysis depicted in figure 30 is again used in the calculation of each quantity in the table. Unless otherwise stated, the diversity phase for all these simulated cases consisted of 2 radians of defocus, Z_4 .

Recall that the total \overline{K} is normalized to unity, and the MSEs given above must be divided by a particular \overline{K} to give a result for a specific light level.

1. Single point source image

It was stated in the introduction that the Cramér-Rao analysis given in this chapter is a generalization of the expressions found in ref. (18). The analysis shown in that paper addressed only the problem of a single point source image, from which $\vec{\alpha}$ estimates were to be extracted, the so-called phase retrieval problem. For the purposes of comparison to the cases above and below, this single-image average CRLB is calculated for the experimental ensemble of 50 quarter-wave phase screens. This quantity was also calculated using equation 101, but with the summation over the index n constrained to a single, in-focus image, $n = 1$. This situation thus represents an estimate of the average minimum error performance for the phase retrieval or blind deconvolution problems discussed in chapter 3 and ref (18), for this class of phase screens. The average error for this case is 8% higher than the best standard phase diversity experiment, shown in item 8 of table 2.

2 and 3. Non-symmetrical beamsplitting

All of the phase diversity situations modelled throughout chapters 3 and 4 of this dissertation have incorporated the assumption that the beamsplitter divides the light equally among the two diversity images. Is this necessarily the optimal procedure? The defocused image of a point source is generally more spatially extended, or spread out, on the image plane. This general feature implies that, given a symmetrical beamsplitter, a larger detector area is being illuminated with the same number of photons. Would there be a benefit in diverting a greater share of the photons to this defocused, diversity image? This question can be investigated via equation 101, by simply modulating the percentages in the appropriate K_n quantities. The indication

here is that there is no advantage in diverting more light to the defocused image, in fact, just the opposite. The greater the deviation from 50/50 beamsplitting, the worse the performance, regardless of the speculations given above. For instance a 40/60 beamsplit yielded 2% greater MSE than the 50/50 case. A 30/70 beamsplit yielded 5% greater MSE.

4. Three-image phase diversity case

Another interesting practical question is addressed by item 4 of table 2: given some finite \overline{K} light level, would there be an advantage in splitting that light equally into three diversity images, as opposed to two? In this demonstration, the evaluation of equation 101 is altered to accommodate three images:

- an in-focus point source image,
- a defocused image with 2 radians of Z_4 , and
- another defocused image with 3 radians of Z_4 .

The normalized light level is split equally, with one-third of the light going to each image. Of course, there is a large number of permutations of defocus and beamsplitting arrangements to choose from. This particular demonstration reveals that, compared to the 2 radian defocused case of figure 31 (item 8), adding the third image resulted in 7% greater MSE. Of course, if a third image path also implied the inclusion of *extra* photons, this MSE would be reduced accordingly, as per equation 101.

5. Two-image extended object phase diversity imaging

One of the advantages touted for the derivation of equation 101 was the ability to account for the imaging of extended objects. In chapters 2 and 4, the assertion was made that phase estimation from point source images acts as a WFS performance upper bound, as compared to the use of images of extended objects (18).

The results of this demonstration bear this out. The gray-scaled data shown in negative in figure 34 represent a CAD rendering of a space object, a Russian Ocean Reconnaissance satellite, which served here as the target object. These data are used in equation 101 for the demonstration of experiment 5 above, to model the pristine extended object irradiance distribution that is convolved with the simulated PSFs.

The Cramér-Rao MSE is seen to be almost two orders of magnitude worse for the imaging of this particular extended object, for the given set of random pupils. One conclusion drawn from this result is that images of extended objects would need to hundreds of times brighter than a corresponding point source to achieve the same theoretical minimum MSE performance.

This result can be understood intuitively by inspecting equation 100. This expression shows that the Fisher information matrix depends on the rate of change of the noiseless image λ with respect to an aberration α_i . The more "sensitive" λ is to changes in α_i , the lower the CRLB, roughly speaking. The image λ consists of either the PSF h for point source imaging, or the object convolved with the PSF, $o * h$, for extended object imaging. Now, h by itself is a direct function of aberrations α_i , through Fourier optics. On the other hand, the object irradiance distribution o is completely independent of the aberrations. Also recall that these images are photo-normalized to remove explicit \bar{K} dependence. Thus a point source image h would reasonably be expected to be more sensitive to changes in α_i than an image which is being convolved with some aberrations-independent o quantity. In extended-source imaging, the aberrations dependent quantity h is being "blurred" by an object o which does not depend on the aberrations at all.

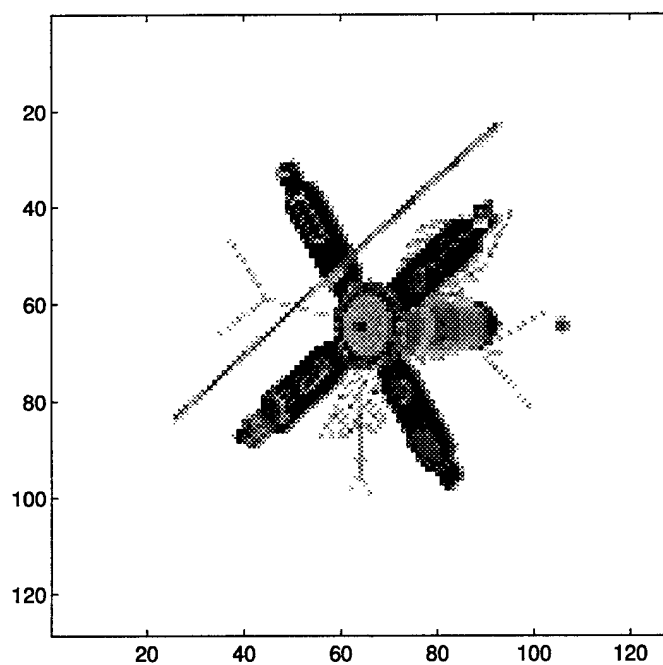


Figure 34. CAD rendering of a Russian Ocean Reconnaissance satellite. The data shown are used to model the pristine "object" irradiance distribution, for imaging of a space object.

6. CRLB in the presence of cone filtering

It was stated in chapter 3 that the noise suppression filtering—used throughout all phase diversity implementations in this research effort—could also be interpreted as the image-domain convolution of a so-called smoothing kernel with the original, unfiltered data, as per ref. (76). Whether noise suppression is interpreted as an image domain convolution, or a Fourier domain lowpass filtering, one obvious disadvantage of the technique is that valid image information is inevitably discarded, along with the noise. Using the various convolutions of equation 101, this experiment demonstrates the Cramér-Rao impact of noise suppression filtering on the same series of 50 quarter-wave pupils used previously. The inverse FFT of the square root of the cone filter shown in figure 10 can be treated as the “object”, similar to the satellite object in experiment 5 above. Remarkably, this allows the effect of noise suppression filtering to be quantified in a fundamental manner. For instance, the demonstration of figure 6 shows that the cone filtering causes the phase estimation MSE to approximately triple. This comparison is with respect to item 8 of table 2, where cone filtering was not accounted for.

7. CRLB for case 1a of chapter 4

The ability to account for cone filtering in a Cramér-Rao sense has been demonstrated in experiment 6 above. Recall that this type of cone filtering was used throughout the simulations phase diversity estimation seen in chapter 4. Recall also that experiment 1a of chapter 4 utilized an ensemble of 50 Zernike aberration phase screens with an average aberration strength of one tenth of a wave. Thus experiment 6 can be replicated to find the CRLB for estimation experiment 1a of chapter 4. Item 7 of table 2 can be used to compare the *actual* estimation performance seen in chapter 4, figure 18, with the theoretical minimum RMS error, by dividing the

item 7 quantity by $\overline{K} = 10^3, 10^4$, and 10^5 photons and taking the square root. Note that the quantities in chapter 4 were given in units of waves, or 2π radians. The minimum *square root* MSE (MRMSE) from the experiment 7 data above are thus compared to the actual root-mean-squared error (RMSE) from chapter 4:

- $\overline{K} = 10^3$; $MRSME = 0.012\text{waves}$; Chapter 4 actual RMSE = 0.018waves
- $\overline{K} = 10^4$; $MRSME = 0.004\text{waves}$; Chapter 4 actual RMSE = 0.005waves
- $\overline{K} = 10^5$; $MRSME = 0.001\text{waves}$; Chapter 4 actual RMSE = $.0.002\text{waves}$

The actual chapter 4 RMSE values, from figure figure 18, approach their theoretical lower bounds to within a factor of 2, when those lower bounds are calculated to account for the information discarded by the cone filtering.

5.5 Conclusion

Previous expressions (18, 62, 76) regarding Cramér-Rao bounds on point-source phase retrieval were slightly generalized in this chapter, in order to account for the multiple images of a phase diversity (26) data set. The fact that phase diversity aberration sensing is not restricted to point-source beacon images is also accounted for in these expressions for the phase diversity Fisher information matrix. The CRLB ends up depending upon the actual, underlying value of the aberrations being estimated. Therefore a numerical, Monte-Carlo analysis of Cramér-Rao quantities was justified. Examples were given of the various types of estimation-theoretic analysis that can be addressed using this type of analysis, questions that have not been analyzed before in an estimation-theoretic sense.

Lacking a specific aberration sensing problem to be evaluated in a CRLB sense, the next best approach appears to be a Monte-Carlo evaluation of a number of pupils of a specific class. Some interesting conclusions can be provisionally inferred from

the particular numerical examples given here. Strictly speaking, the conclusions are only valid for the set of pupils used in the Monte-Carlo evaluations, but the results could point the way towards further research. For instance, there may be an advantage in performing phase diversity estimation on symmetrically defocused image pairs instead of the more traditional image set.

Similarly, for the Monte-Carlo ensemble of quarter-wave aberrated pupils discussed above, it appeared that a diversity defocus of 2 radians was measurably better in terms of minimum MSE than, say, a defocus of 3 radians. There also did not appear to be any advantage to splitting a finite amount of light among 3 diversity images as opposed to only 2 for this aberration set. The assertion that extended objects are less advantageous phase diversity imaging targets than point sources was verified.

It is important to stress that all of these conclusions are based on CRLB calculations using the same ensemble of 50 aberration phase screens. In the strictest sense, these conclusions apply only to those 50 pupils and no others. Whether or not 50 screens is sufficient to allow meaningful generalization of these results remains an open question for future research. It may well be that many more than 50 pupils are needed to give results which can be generalized to an entire statistical class of pupils. Also, the numerical results shown here all involve theoretical minimum error levels which may never be attained by practical, realizable estimators. Finally, since these results were meant only for purposes of demonstration, it is important to realize that the aberration setup is highly contrived. It is unlikely that an optical situation could be found where only Zernike modes 4 through 11 could be found, in equal average RMS strengths.

The actual PDWFS average estimation errors for case 1a of chapter 4 were compared to their corresponding Cramér-Rao limits. All three cases were found to be within a factor of 2 or closer to their CRLBs. These generalized Cramér-Rao expressions also provide a tool for quantifying the estimation-theoretic impact of noise suppression filtering. None of these general, fundamental questions have been addressed before in the the existing body of phase diversity literature.

VI. Applications of phase diversity to ground-based adaptive optical systems: Diagnosing image-path-only aberrations

6.1 Problem introduction

Adaptive optics (AO) systems are designed to correct the dynamically changing optical phase delays introduced by the Earth's turbulent atmosphere. Under certain conditions, these systems can approach nearly diffraction limited imaging performance (21, 72, 77). The presence of a non-common optical aberration that is present only in the AO imaging path, as depicted in simplified, schematic form in figure 35, can result in a serious degradation of this performance. Clearly, a fixed image-path aberration like this would not be detected by the wavefront sensor (WFS) of the AO system. Therefore, there is no way for the AO system to apply appropriate correction to the deformable mirror (DM). One way of prescribing such an aberration would involve an actual hands-on diagnosis of the system hardware in the optical path of the imaging camera, an approach which may not be feasible given cost and operational constraints. Short of this manpower intensive approach, it would be advantageous to diagnose such an aberration using the information that is already available, namely, the image data and the system WFS measurements of the phase residual after the DM. This chapter outlines a novel extension and application of the *phase diversity* technique, applied to the problem of estimating this unsensed aberration. In implementing this method, outlined below, it will be shown that it is possible to diagnose this aberration using a collected ensemble of image data and corresponding WFS residual phase measurements. The initial inspiration for this novel application of phase diversity theory is due to Brent Ellerbroek, a scientist at the USAF Phillips Laboratory Starfire Optical Range.

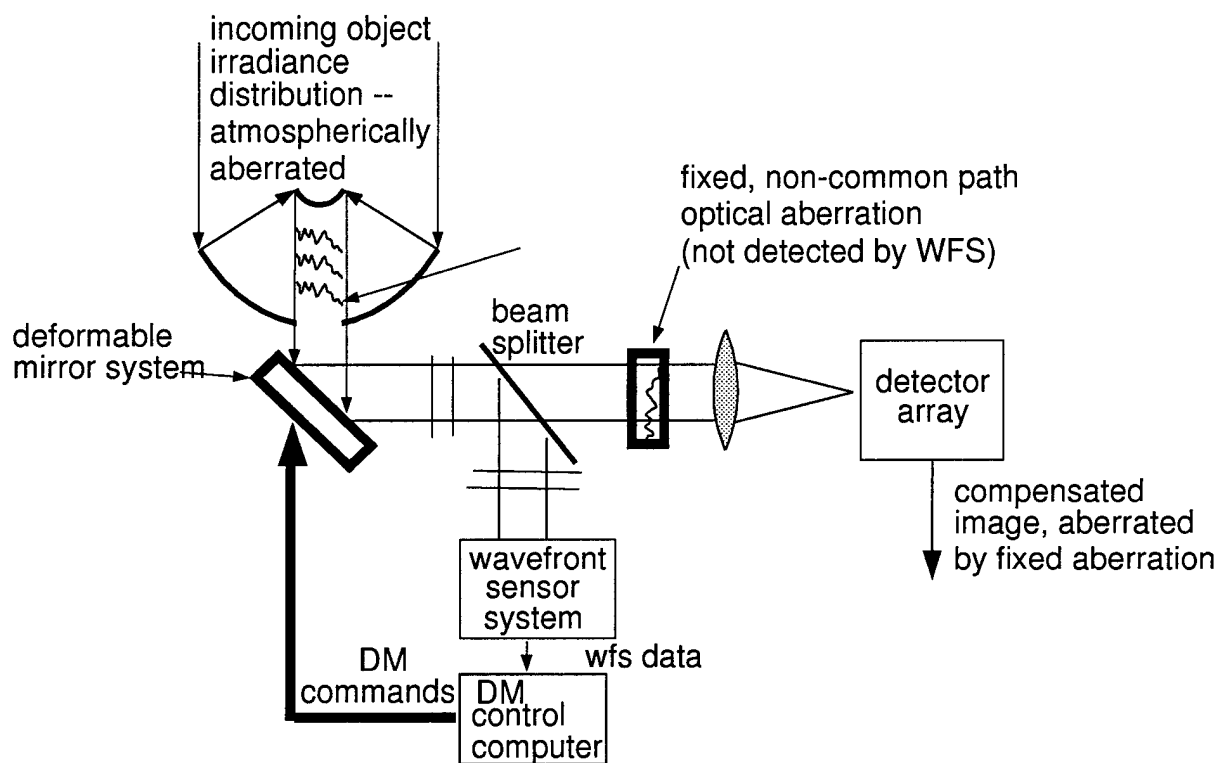


Figure 35. Simplified block diagram of an adaptive optics imaging system, depicting an aberration that is present in the optical path that the imaging camera does not share with the WFS.

The remainder of this chapter is organized as follows. First the phase diversity technique, which has been discussed and analyzed extensively in preceding chapters, is re-interpreted such that it can be adapted to the problem of diagnosing the aberration of figure 35. Then an actual example of this extended aberration diagnosis method is presented, using data from an operational adaptive optics system: the USAF Phillips Laboratory Starfire Optical Range system.

6.2 *Adapting phase diversity to this problem*

The general PDWFS method involves the collection of multiple images of the same target, each image due to a slightly different pupil phase. Throughout previous chapters of this dissertation, these diversity phases have been simulated as varying amounts of pupil defocus, because defocus is relatively easy to implement in optical hardware for an intentional phase diversity system, such as the concept shown in figure 1 in chapter 1. This traditional PDWFS formulation will later be contrasted with a so-called unintentional phase diversity setup. The required PDWFS data set consists of N noisy, phase diverse images, along with the corresponding knowledge of how the various pupil phases differ from each other. This data set was written symbolically in chapter 3 as:

$$\begin{aligned} \{d_n(\vec{x}; \vec{\alpha}_n)\} &= \{d_1(\vec{x}; \vec{\alpha} + \vec{\Delta}_1), d_2(\vec{x}; \vec{\alpha} + \vec{\Delta}_2), \\ (103) \quad &\dots, d_N(\vec{x}; \vec{\alpha} + \vec{\Delta}_N)\}, \end{aligned}$$

where the d_n quantities still represent the n th noisy collected image, while the $\vec{\alpha}_n$ vectors represent aberration coefficients, which is the resultant of two different aberration components: $\vec{\alpha}_n = \vec{\alpha} + \vec{\Delta}_n$. The fixed aberration to be estimated is designated

by the coefficient vector $\vec{\alpha}$, and the *known* phase diversity aberrations are symbolized by the coefficient vectors $\vec{\Delta}_n$.

For the purposes of this chapter, the most important fact to recall from this summary of the phase diversity principle is that the diversity phase screens are in no way restricted to the class of parabolic defocus aberrations. Also, the number of images is not restricted to only two. In order to use the Gonsalves technique to measure the fixed aberration depicted in figure 35, all that is needed is a set of images derived from different pupil phase screens, and knowledge of just how those phase screens differ from each other. Ideally, the images and WFS frames from an AO system will provide such a generalized, unintentional phase diversity data set, as discussed below.

But recall the underlying problem depicted in figure 35, namely the presence of an unknown aberration in the imaging path of an AO system. This aberration means that the WFS phase estimates provided by the WFS system are all incorrect. But, assuming that this unknown aberration is temporally quasi-static, the WFS phase data are all incorrect by some fixed unknown additive constant. Therefore, the *difference* between any two phase estimates is unaffected by this unknown aberration. Thus the difference between any two phase screens gives information regarding the diversity phase between the corresponding images. These diversities are the $\vec{\Delta}_n$, referenced above in equation 103.

For the ground-based AO case of figure 35 the required, unintentional, phase diversities arise from the temporal differences in residual pupil phase errors from one exposure to the next. These temporal differences are due to the dynamic aberrations caused by the turbulent atmosphere, combined with the imperfect compensation of the AO system. Therefore an AO image ensemble and the corresponding WFS

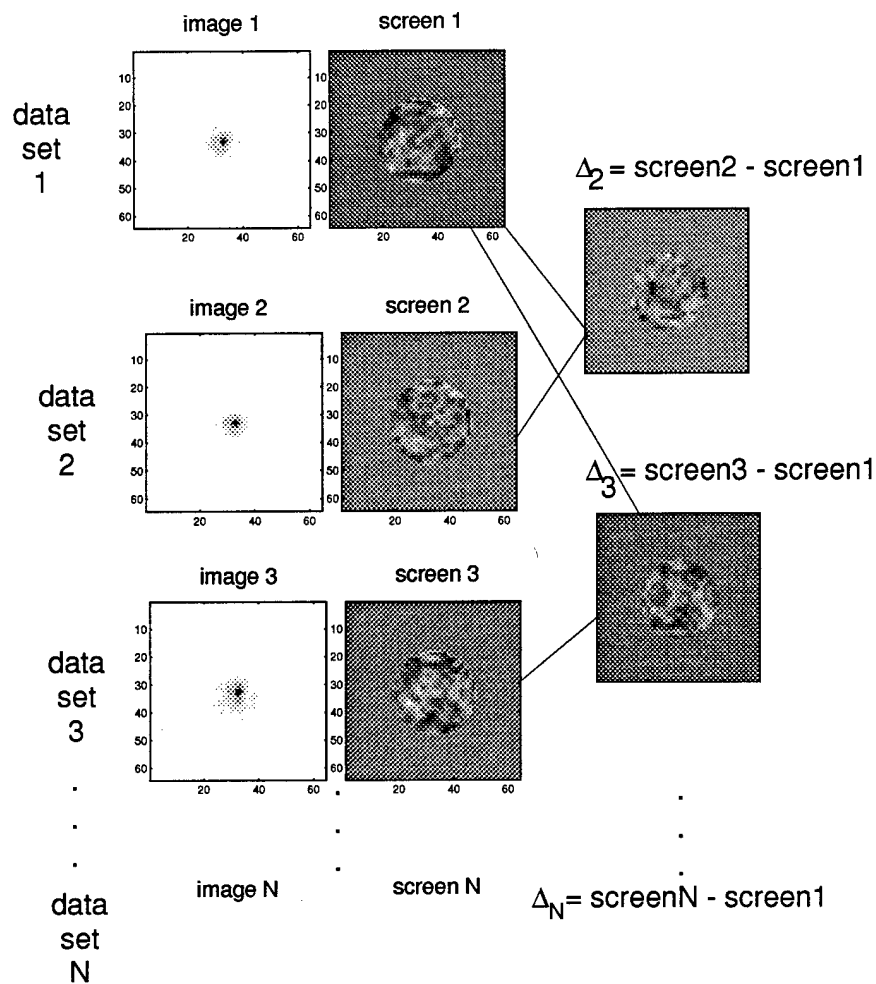


Figure 36. Schematic for using phase diversity to diagnose an AO image aberration that is not present in the WFS optics path, by relying only on the differences in WFS measurements. The differences are due to the dynamic action of the atmosphere-adaptive optics system.

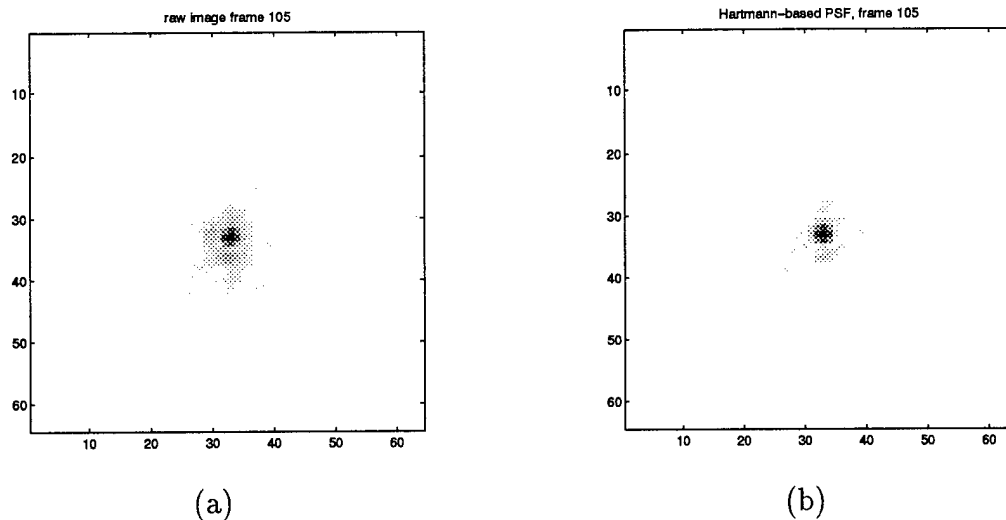


Figure 37. Example images frames from ensemble of SOR image and WFS data. Image-based PSF estimate is shown in (a), WFS-based PSF estimate is shown in (b). Images shown in negative for clarity.

measurements of residual phase errors constitute a generalized phase diversity data set, consisting of a collection of N phase diversity image frames, as well as knowledge of the corresponding non-parabolic diversity phases. The diversity information is obtained through simple subtraction of the appropriate WFS-derived phase screen data. This concept is depicted graphically in figure 36. With this approach a phase diversity data set can be assembled without intentionally creating a phase diversity system like that shown in figure 1. The traditional phase diversity concept is now extended to encompass more than two total images, and non-quadratic diversity phases that are now due to the temporal differences between subsequent residual AO phase screens.

6.3 Example application

In the course of an experiment in *a posteriori* WFS-based deconvolution of AO imagery as in ref. (70), the U. S. Air Force Phillips Laboratory Starfire Optical

Range (SOR) provided this institution (AFIT imaging research consortium) with adaptive optics (AO) compensated image ensembles, where the target object was the star Harvard Reference no. 2286 listed in ref (35). The SOR also provided an ensemble of corresponding Hartmann WFS (H-WFS) slope measurements. The data were collected using the 1.5 meter AO system described in reference (21). With these sources of data, two different types of estimates of the system point spread functions (PSFs) were available:

1. PSFs as shown in the star images; and
2. PSFs as derived from the WFS data.

In comparing the two types of PSF estimates for our data set, it became apparent that the image-PSFs exhibited an aberration that was not present in the Hartmann-PSFs.

The presence of this image-path aberration is demonstrated in the various types of PSF estimates shown in figure 37. Each PSF estimate shows the 64×64 pixels of an original 64×64 pixel image, with negative images shown for clarity. The imaging setup for the telescope used to gather the data is such that the images were Nyquist sampled, 2 pixels per λ/D . The two types of PSF estimates are shown side-by-side in figure 37 (a) and (b) for one typical image/WFS data frame pair. Figure 37 (a) shows the image detected by the camera

$$(104) \quad PSF_{im} = d_i(\vec{x}),$$

while figure 37 (b) shows the PSF estimate obtained by the WFS

$$(105) \quad PSF_{wfs} = \left| \text{FT} \left[W \exp[j\phi_i] \right] \right|^2,$$

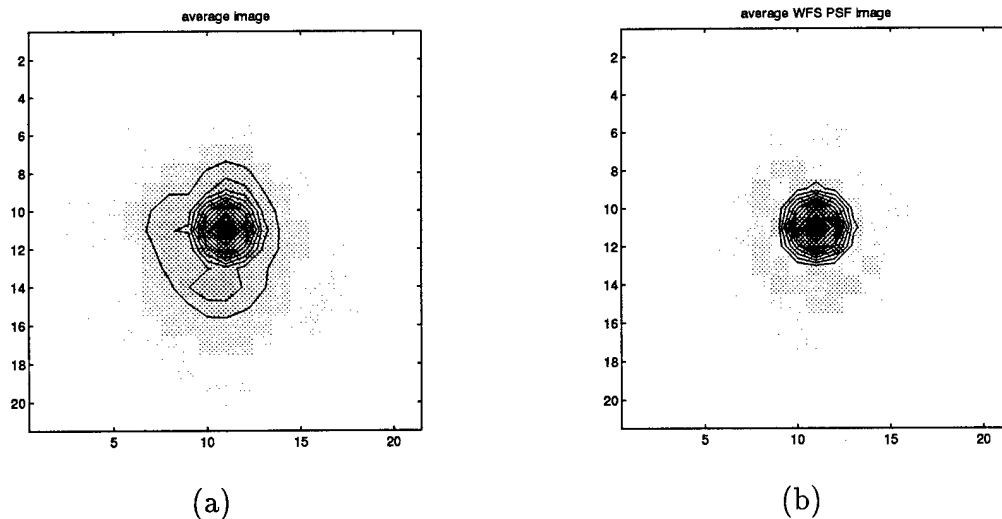


Figure 38. 50-element ensemble averages of PSF estimates: star image frame (a) and its corresponding original Hartmann WFS-based PSF estimate (b). These average estimates show how the Hartmann is not detecting the unknown aberration.

where W is the aperture transmittance function of the telescope, and ϕ_i is the i th WFS estimate of the residual phase.

An asymmetrical lobe feature extends from the lower center-left of the main PSF body shown figure 37(a). This feature clearly does not appear in the WFS-based PSF estimate of figure 37(b). The possibility that the lobe is caused by a companion star has been ruled out — the Yale Bright Star Catalog (35) lists HR2286 (13μ Geminorum) as a 3 star system, but the secondary and tertiary stars are not within the field of view in these image ensembles.

Figures 38(a) and (b), show that this aberration is not isolated to a single image, but rather is exhibited across the image ensemble of this particular observation run. Where figure 37(a) and (b) show images derived from a single data realization, figure 38(a) and (b) show 50-image ensemble averages of the same quantities, now zoomed in on the central 21×21 pixels, with contour lines overlayed. Figure 38(a)

depicts the quantity

$$(106) \quad \langle PSF_{im} \rangle = \frac{1}{50} \sum_{i=1}^{50} d_i(\tilde{x}),$$

while figure 38(b) shows

$$(107) \quad \langle PSF_{wfs} \rangle = \frac{1}{50} \sum_{i=1}^{50} \left\{ \left| \text{FT} \left[W \exp[j\phi_i] \right] \right|^2 \right\},$$

These images imply that the image-path aberration is apparent throughout the image ensemble, but the WFS consistently fails to detect this aberration, bolstering the hypothesis of figure 35.

This hypothesis is also demonstrated in the frequency domain by figure 39. The lower 'o' curve represents the modulus of the radially-averaged profile of the Fourier transform of the average, image based PSF:

$$(108) \quad \text{curve}_o = \left| \text{Radial average} \left(\text{FT} \left[\langle PSF_{im} \rangle \right] \right) \right|.$$

The upper '+' curve shows the same for the Hartmann WFS based average OTF estimate:

$$(109) \quad \text{curve}_+ = \left| \text{Radial average} \left(\text{FT} \left[\langle PSF_{wfs} \rangle \right] \right) \right|.$$

In order to estimate the image path aberration, the noise suppressed Gonsalves algorithm, using $N = 10$ sets of image/WFS data, was implemented. Before executing the algorithm, the two-dimensional Fourier transform image spectra were digitally filtered using a circularly symmetric, linearly tapered cone filter, which acted to include, but linearly attenuate, the contribution of image data spatial fre-

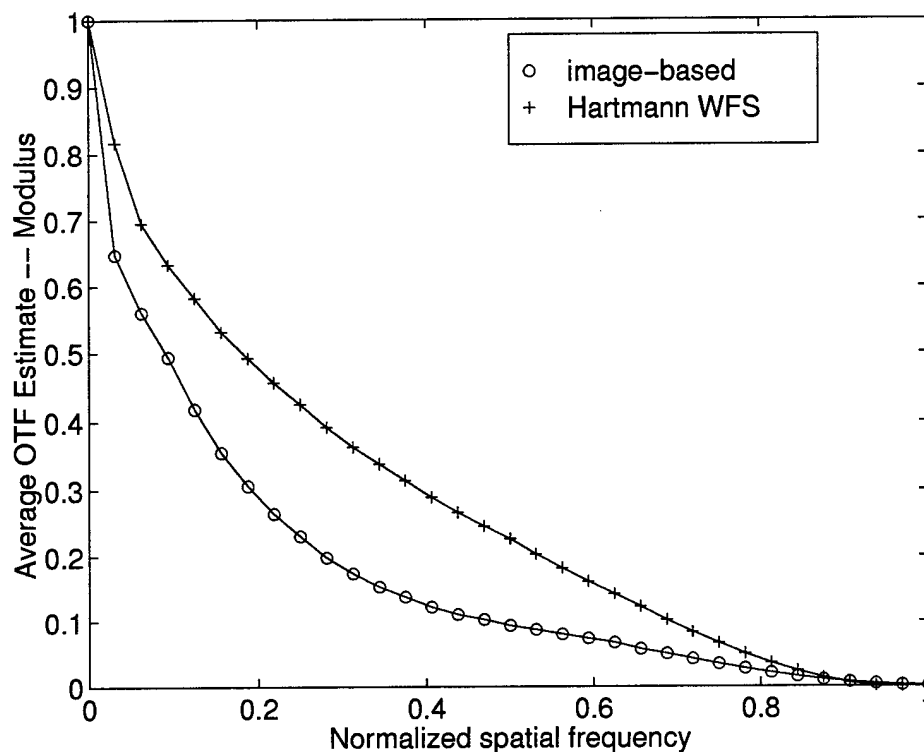


Figure 39. The lower curve represents the modulus of the radially-averaged profile of the Fourier transform of the average, image based PSF. The upper curve represents the same, but for the Hartmann-WFS derived PSF estimate.

quencies up to one-half of the diffraction-limited spatial frequency, and excluded all higher frequency image information. This suppression of upper frequency noise is required in order stabilize the iterative minimization procedure and allow it to reliably converge, as discussed in chapter 3 (12, 49).

The Gonsalves function was then minimized by searching, via a quasi-Newton method (30, 64) over the space of Zernike polynomials 4 through 22, using the Zernike polynomial conventions specified by Noll (58).

The Zernike coefficients that minimize the Gonsalves objective function ideally represent the total aberration that is manifest on the image plane. Therefore, the unknown fixed image-path aberration can be obtained by subtracting the Hartmann WFS aberration estimate from the phase diversity estimate:

$$(110) \quad \vec{\alpha}_{\text{unknown}} = \vec{\alpha}_{\text{pd}} - \vec{\alpha}_{\text{Hartmann}},$$

where the $\vec{\alpha}$ are now 19-element vectors representing the Zernike coefficients for modes 4-22 of the appropriate phase screen. These Zernike coefficients for the diagnosed aberration are presented in table 3, in units of radians. The table shows that the strongest contributors in this case was mode 10, corresponding to a triangular coma, or clover aberration. The phase screen represented by the 8 coefficient estimate is shown in table 3 is shown in figure 40 in gray scale, where gray level value corresponds to optical path-length difference across the pupil.

Figure 41 represents the so-called "augmented" WFS estimate of the average PSF, where the phase diversity estimate of the aberration has been added to the

Table 3. Noll-modified Zernike coefficients of the diagnosed image-path aberration phase screen, 19-mode estimate.

Noll-Zernike mode no.	coefficient (radians)
4	0.0036
5	0.1310
6	-0.2123
7	0.2182
8	-0.0532
9	-0.0814
10	-0.3567
11	0.2461
12	0.0326
13	0.0797
14	-0.0409
15	0.1625
16	0.0829
17	-0.0465
18	0.1049
19	-0.2705
20	-0.2986
21	0.0951
22	0.0103

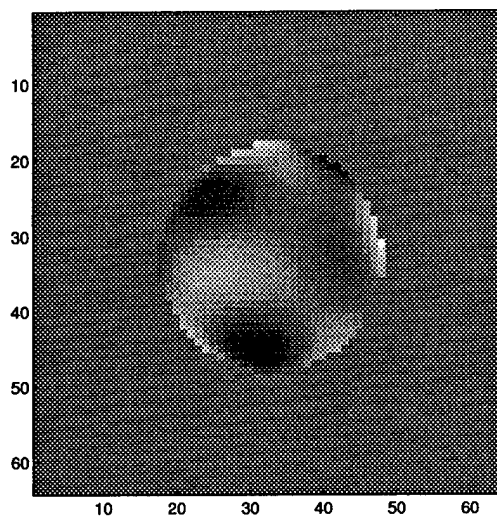


Figure 40. The phase diversity estimate of the image-path aberration, restricted to Zernike modes 4-22. Gray-level map is shown.

H-WFS phase estimates:

$$(111) \quad \langle PSF_{aug} \rangle = \frac{1}{50} \sum_{i=1}^{50} \left\{ \left| \text{FT} \left[W \exp[j(\phi_i + \phi_{pd})] \right] \right|^2 \right\},$$

In comparing figure 41 to figure 38(a), the original average image, it appears that the the lobe feature extending from the lower central left of the main PSF in the image data has been somewhat recaptured by these augmented WFS data PSF estimates. The augmented estimate is not perfect, due probably to the limited number of aberration degrees-of-freedom estimated here. But there is good reason to believe that the significant fraction of the image path aberration has been estimated reasonably well in this case.

Correspondingly, in the frequency domain, figure 42 shows that the phase diversity augmented average OTF estimate matches that for the image-based OTF estimate fairly well. The lower 'o' curve represents the modulus of the radially-

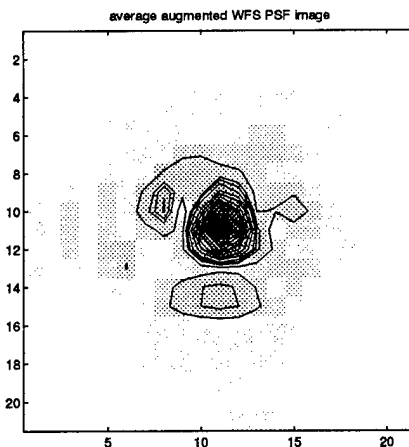


Figure 41. Average phase-diversity-augmented, WFS-based PSF estimate, 19-mode case, as discussed in the text. Again, central 21 pixel squares are shown, in negative. Compare to figure 38(a), the original average image.

averaged profile of the Fourier transform of the average, image-based PSF, as before in figure 39. And the upper '+' curve represents the same, but for the Hartmann-WFS derived PSF estimate. The third '*' curve corresponds to the phase-diversity augmented average OTF,

$$(112) \quad \text{curve}_* = \left| \text{Radial average} \left(\text{FT} \left[\langle PSF_{aug} \rangle \right] \right) \right|,$$

showing how the overly optimistic H-WFS estimate has been driven back towards the image-based estimate.

6.4 Remarks on practical application of the technique

This novel technique represents a potentially powerful and valuable tool that is available to the practicing astronomer who is using an operational adaptive optical system. Adaptive optical systems are complex, expensive pieces of hardware, de-

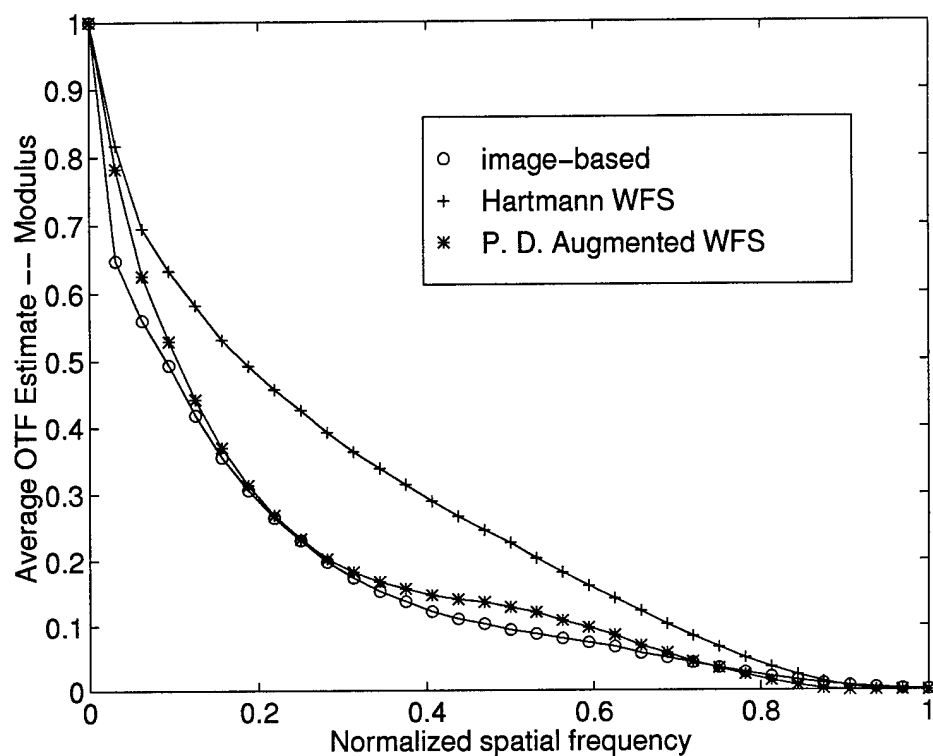


Figure 42. The lower 'o' curve represents the modulus of the radially-averaged profile of the Fourier transform of the average, image-based PSF. The upper '+' curve represents the same, but for the Hartmann-WFS derived PSF estimate. The third '*' curve corresponds to the phase-diversity augmented average OTF, showing how the overly optimistic H-WFS estimate has been driven back towards the image-based estimate.

signed with the goal of attaining as close to diffraction-limited imaging performance as possible. If the effect of an image path aberration is later found to be present in a collection of imagery, as happened in the experimental case shown here, the hard-won imaging and resolution performance gains attained by the expensive AO system are at least partially negated. This chapter has shown a procedure for diagnosing that error, and possibly regaining the lost imaging performance.

Assume, for example, that the cause of some fixed image-path aberration cannot be tracked down, or that such an aberration is an unavoidable side effect of some piece of science hardware on the telescope optics bench. The estimates obtained from the technique shown in this chapter could then be used for pre-emptive compensation of the aberration, in the following manner. The experimenters would first collect a series of calibration images, which, due to the flexibility of the Gonçalves phase diversity scheme, need not be point source images. Next, they would implement the procedure demonstrated in this chapter to estimate the unavoidable image-path aberration. Then the appropriate conjugate of this aberration estimate could be imposed as a fixed bias upon the AO deformable mirror. When collecting subsequent science images, the dynamic deformations needed for atmospheric compensation would be added to this fixed bias deformation by the AO system. This fixed deformation would act to pre-correct for the image-path aberration throughout the observing session.

Alternatively, recall again, as just stated, that the phase diversity technique does not explicitly require point source imagery (62). One could therefore also use these fixed aberration estimates *ex post facto*, in a post-processing scheme. The procedure would be similar to the general DWFS scheme presented in chapter 2 or ref. (70). But now the WFS estimate would be augmented by the fixed aberration

estimate. One could then deconvolve the effects of a fixed image-path aberration on extended object imagery that has already been collected, assuming that the corresponding WFS data were also recorded during the observing session.

6.5 Conclusion

This chapter has outlined a potential technique for diagnosing non-common image-path aberrations in an AO system that are not sensed by the WFS. This technique was demonstrated using actual astronomical imagery from an operational adaptive optics system, along with the corresponding wavefront sensor measurements. The astronomical imagery appeared to exhibit an aberration that was not detected by the WFS system. Using the dynamic, temporal differences between WFS estimates as diversity phases, the Gonsalves technique was adapted to this novel situation, and estimates of this unsensed aberration were obtained and demonstrated.

Presumably, an even more accurate estimate could have been obtained in the examples shown here, through refinement of the algorithm, which was not attempted. Estimation of even more Zernike coefficients might be found to be feasible by

- optimizing the number N of images used,
- refining the noise suppression method (12, 49), or
- experimenting with the more complicated, object dependent Poisson maximum-likelihood technique discussed in reference (62).

It is worthwhile to note that the straightforward, approach shown here yielded a reasonable aberration estimate even before such optimization or experimentation.

Finally, this chapter represents another important validation of the phase diversity concept, made all the more significant in the use of actual imagery obtained

from an operational telescope. This fact stands in contrast with the numerically simulated imagery of the type used in previous chapters, and used in a large fraction of the existing phase diversity literature.

VII. Conclusion

This chapter briefly summarizes the outcomes of this research effort, and presents ideas for follow-on research problems extending from the ideas studied here.

This dissertation has presented an in-depth analysis of the phase diversity wavefront sensing (PDWFS) technique, applied to the special problem of astronomical imaging through space-based telescopes. The overall PDWFS problem is concerned with the estimation of pupil aberrations from pairs of in-focus and defocused images. The aberration estimation accuracy of the technique was investigated in terms of numerical, Monte-Carlo computer imaging experiments, incorporating photon noise, as well as in terms of fundamental, estimation-theoretic performance limits. Neither of these types of analysis have been applied to this problem in the existing published literature.

7.1 Summary of results

Feasibility of general, WFS-based deconvolution, as demonstrated in chapter 2, along with the experimental and theoretical performance of PDWFS, lead to the conclusion that the integration of phase diversity into the operational design of a space telescope would be advantageous. The wavefront estimation accuracies achieved here are sufficient to cause an improvement in imaging performance whether the estimates are used to:

1. mechanically correct the aberrations using mirror phasing actuators, or
2. mathematically deconvolve the aberration effects in post-processing.

The conclusion holds regardless of whether a monolithic mirror or a segmented mirror is used in the telescope. But PDWFS is especially relevant for a segmented mirror,

since the discontinuous nature of an aberrated phase screen, caused by misaligned segments, is not amenable to diagnosis by a standard slope-sensing WFS.

Numerical experimentation on the PDWFS technique for the case of space telescope point-source imaging showed that accurate pupil estimates can be obtained even under very low light conditions. By way of illustration, in one set of experiments on estimating the 0.10λ RMS petal piston errors of a segmented space telescope, using very dim, photon-limited, point-source images with $\overline{K} = 1000$, phase estimation RMS errors of 0.012λ were found. The segmented mirror experiments here incorporated a model of the Next Generation Space Telescope (NGST) proposal, an idea for a possible follow-on to the Hubble Space Telescope.

Some of the photon noise limitations of the technique were also encountered. In one example case, estimates similar to those in the previous paragraph were made, but *both* segment piston and tilt misalignment errors were included. Under the same low-light conditions, 20% of these higher degree-of-freedom Monte-Carlo PDWFS experimental cases ended in failure, the algorithm converging on WFS estimates that were incorrect by orders of magnitude. These failure outcomes are consistent with the fact that the Gonsalves PDWFS estimator is not optimized to photon statistics. This appears to allow the least-squares error metric search to become trapped in invalid minima which are due to noise instead of actual aberrations.

The Cramér-Rao lower bound analysis presented here represents a fundamental new approach towards assessing the advantages and disadvantages of any particular phase diversity imaging configuration. One key result of such an analysis, demonstrated in chapter 5, is the proposal that phase diversity WFS estimation might be more appropriately carried out using symmetrically defocused images. Another finding is that the optimal phase diversity defocus for a given set of quarter wave

aberrations appears to be approximately 2 radians. The optimal amount of defocus to use in phase diversity imaging is a fundamental practical question that has not been addressed before in any fundamental way. A variety of other modifications to the standard PDWFS implementation are proposed and investigated via CRLB analysis, such as alternate forms of beamsplitting and target objects. The performance of PDWFS Monte-Carlo experiments was also analyzed in terms of this theoretical limit.

In the course of this research, a novel method for prescribing adaptive optics aberrations was also addressed. A potentially useful new tool has been made available to the AO astronomical imaging community through this novel re-interpretation of the phase diversity algorithm. The technique was demonstrated using actual imagery from an operational AO system, with successful results. The first 22 modes of a troublesome image-path aberration appear to have been accurately captured by this technique.

7.2 Follow-on research ideas

Ideas for follow-on research include the interesting possibility of using a weighted least-squares methodology for including model information into the phase diversity technique. Preliminary calculations show that the inverse covariance matrix of the Fourier transform of a photon limited image spectrum, which incorporates knowledge of the object and optical transfer function, could serve as a least-squares weighting matrix.

Other research projects include numerical simulation of other telescope aberration models, besides the examples given here. These investigations could, of course, be expanded beyond the realm of space-based telescopes. Sparse, phased array tele-

scopes would be good candidates for phase diversity misalignment sensing (61), for example.

Similarly, the Cramér-Rao numerical evaluations given in chapter 5 could be replicated for a large variety of phase diversity configurations and experiments not evaluated here. Sparse aperture systems and adaptive optical telescopes along the lines of the application in chapter 6 could all be analyzed under the CRLB paradigm, possibly allowing similar, fundamental types of questions to be addressed. CRLB analysis could also be used to determine whether, for instance NGST tilt parameters are more difficult to estimate than similar piston parameters.

The uniqueness and ambiguity example shown in chapter 3 is also amenable to further study. One idea for quantifying the concepts embodied in the ambiguity demonstration of chapter 3 is to use a numerical, Monte-Carlo analysis, along the lines of the the analogous phase retrieval uniqueness study given in ref. (75). For instance, the single-image phase retrieval case could be compared with the phase diversity case, in order to determine if there is some way to quantify the improved ambiguity properties of the phase diversity problem, in light of the fact that PDWFS is automatically constrained by the collection of multiple, diverse images.

Another straightforward problem waiting to be tackled involves quantifying the performance of the Poisson maximum-likelihood formulation of phase diversity, along the lines of the Monte-Carlo study of chapter 3. The object-dependence issue could be sidestepped for a preliminary study, wherein it could be assumed that the object distribution were known ahead of time, such as point-source imaging of a known single star, in order to diagnose pupil aberrations. Under this framework, the object dependence would cease to be a problem.

Appendix A. Remarks on pupil weighting and PSF normalization

The desired end result of the pupil "pre-normalization" mentioned in the derivation of chapter 5, the Cramér-Rao chapter, is

$$(113) \quad \int_{\vec{x}} h(\vec{x}; \vec{\alpha}) d\vec{x} = 1.$$

From Fourier identities, and the derivations shown in section 5.2.1 we therefore can work backwards to obtain

$$(114) \quad \int_{\vec{x}} h(\vec{x}; \vec{\alpha}) d\vec{x} = 1$$

$$(115) \quad \int_{\vec{x}} |G(\vec{x}; \vec{\alpha})|^2 d\vec{x} = 1$$

$$(116) \quad \int_{\vec{u}} |g(\vec{u}; \vec{\alpha})|^2 d\vec{u} = 1$$

$$(117) \quad \int_{\vec{u}} g(\vec{u}; \vec{\alpha}) g^*(\vec{u}; \vec{\alpha}) d\vec{u} = 1$$

$$(118) \quad \int_{\vec{u}} W(\vec{u}) e^{j\phi(\vec{u}; \vec{\alpha})} W(\vec{u}) e^{-j\phi(\vec{u}; \vec{\alpha})} d\vec{u} = 1$$

$$(119) \quad \int_{\vec{u}} W^2(\vec{u}) d\vec{u} = 1$$

using, for example, table 9-1 of reference (22).

Bibliography

1. "What will be the next Big Thing," *Nature*, 381:465 (1996).
2. Acton, D. S., et al. "Full-field wavefront measurements with phase diversity," *Astronomy and Astrophysics*, 309:661-672 (1996).
3. Ahmad, A., et al. "Tolerance analysis versus image quality: a case study for cost-effective space optics," *Opt. Eng.*, 34:575-583 (1995).
4. Babcock, H. W. "The possibility of compensating astronomical seeing," *Publ. Astron. Soc. Pac.*, 65:229-236 (1953).
5. Bates, R. H. T. and H. Jiang. "Blind deconvolution — recovering the seemingly irrecoverable!," *International Trends in Optics* edited by J. W. Goodman, Boston: Academic Press, Inc. International Commission for Optics.
6. Born, M. and E. Wolf. *Principles of Optics*. New York: Pergamon, 1964.
7. Burrows, C. J., et al. "Imaging performance of the Hubble Space Telescope," *Astrophys. J.*, 369:L21 (1991).
8. Carreras, R. A., et al. "A laboratory experiment using phase diversity to extract higher order Zernike coefficients," *SPIE Proc.*, 2302:323-329 (1994).
9. Chaisson, E. J. *The Hubble Wars*. New York: Harper Collins, 1994.
10. Chanan, G., et al. "The W. M. Keck telescope phasing camera system," *SPIE Proc.*, 2302:1139-1150 (1994).
11. Chi, C., et al. "Vibrational modes of the primary mirror structure in the Large Space Telescope system," *Space Optics, Proc. of the Ninth International Conference of the International Commission for Optics*, 9:209-238 (1974).
12. Craig, I. J. D. and J. C. Brown. *Inverse Problems in Astronomy: a guide to inversion strategies for remotely sensed data*. Bristol, England: Adam Hilger Ltd., 1986.
13. Dooling, D. "Beyond Hubble," *The Institute, IEEE monthly newsletter*:1 (June 1996).
14. Dressler, A. "HST and Beyond. Available from <http://saturn1.gsfc.nasa.gov/ngst/>," *NASA NGST home page* (Committee report, 1996).
15. Ellerbroek, B. and D. Lee. *Personal conversation*. Starfire Optical Range, Phillips Laboratory, Kirtland AFB, NM.
16. Ellerbroek, B. and D. Morrison. "Linear Methods in phase retrieval," *SPIE Proc.*, 351:90-95 (1982).
17. Fienup, J. R. "Phase-retrieval imaging problems." *International Trends in Optics* edited by J. W. Goodman, Boston: Academic Press, Inc. International Commission for Optics.

18. Fienup, J. R., et al. "Hubble Space Telescope characterized by using phase-retrieval algorithms," *Applied Optics*, 32:1747-1767 (1993).
19. Fried, D. L. "Least-square fitting a wave front distortion estimate to an array of phase-difference measurements," *J. Opt. Soc. Am.*, 67:370-375 (1977).
20. Fried, D. L. "Post-Detection Wavefront Distortion Compensation," *SPIE Proc.*, 828:127-133 (1987).
21. Fugate, R. Q., et al. "Two generations of laser-guide-star adaptive-optics experiments at the Starfire Optical Range," *J. Opt. Soc. Am.-A*, 11:301-314 (1994).
22. Gaskill, J. D. *Linear Systems, Fourier Transforms, and Optics*. New York: John Wiley & Sons, 1978.
23. Gates, E. L., et al. "Phase diversity as an on-line wavefront sensor: experimental results," *SPIE Proc.*, 2302:330-339 (1994).
24. Gonsalves, R. A. "Fundamentals of wavefront sensing by phase retrieval," *SPIE Proc.*, 351:56-65 (1982).
25. Gonsalves, R. A. "Phase retrieval and diversity in adaptive optics," *Opt. Eng.*, 21:829-832 (1982).
26. Gonsalves, R. A. and R. Chidlaw. "Wavefront sensing by phase retrieval," *SPIE Proc.*, 207:32-39 (1979).
27. Gonzalez, R. C. and R. E. Woods. *Digital Image Processing*. Reading, Massachusetts: Addison-Wesley, 1993.
28. Goodman, J. W. *Introduction to Fourier Optics*. New York: McGraw-Hill, 1968.
29. Goodman, J. W. *Statistical Optics*. New York: John Wiley & Sons, 1985.
30. Grace, A. *Optimization Toolbox for use with MATLAB*. The Math Works, Inc., MA, 1992.
31. Hanisch, R. J. "Image restoration for the Hubble Space Telescope," *SPIE Proc.*, 2198:1349-1356 (1993).
32. Hardy, J. W. "Active Optics: A New Technology for the Control of Light." *Proceedings of the IEEE* 66. 651-697. 1978.
33. Hayes, M. H., et al. "Phase-only Signal Reconstruction," *IEEE ICASSP Proc.*, CH1559:436-440 (1980).
34. Hecht, E. *Optics*. Reading, MA: Addison Wesley, 1987.
35. Hoeffleit, D. and C. Jaschek. *The Bright Star Catalogue, 4th revised edition*. New Haven: Yale University Observatory, 1982.
36. Horner, J. L. and P. D. Gianino. "Phase-only matched filtering," *Appl. Opt.*, 23:812-816 (1984).

37. Horton, R. F., et al. "The MSTI-3 sensor package optical design," *SPIE Proc.*, 2478:53-74 (1995).
38. Howell, W. E. "Recent advances in optical control for large space telescopes," *Space Optics, Proc. of the Ninth International Conference of the International Commission for Optics*, 9:239-258 (1974).
39. Jakubowski, A. K., et al. "Eight-meter UV/visible/IR space telescope," *SPIE Proc.*, 2478:20-34 (1995).
40. Jefferies, S. M. and J. C. Christou. "Restoration of astronomical images by iterative blind deconvolution," *Astrophys. J.*, 415:862-874 (1993).
41. Jones, P. A., et al. "Passive primary mirror concept for SOFIA," *SPIE Proc.*, 2478:188-194 (1995).
42. Kay, S. M. *Modern Spectral Estimation*. New Jersey: Prentice Hall, 1988.
43. Kay, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice Hall, 1993.
44. Kittel, C. and H. Kroemer. *Thermal Physics*. New York: W. H. Freeman and Company, 1980.
45. Labeyrie, A. "Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier Analysing Speckle Patterns in Star Images," *Astronomy and Astrophysics*, 6:85-87 (1970).
46. Lane, R. G. "Blind deconvolution of speckled images," *J. Opt. Soc. Am. A*, 9:1508-1514 (1992).
47. Langendjik, R. L., et al. "Maximum likelihood image and blur identification: a unifying approach," *Opt. Eng.*, 29:422-435 (1990).
48. Lee, D. J., et al. "Using wavefront sensor information in image post-processing to improve the resolution of telescopes with small aberrations," *SPIE Proc.*, 2863 (1996 (to be published)).
49. Lofdahl, M. G. and G. B. Scharmer. "Wavefront sensing and image restoration from focused and defocused solar images," *Astronomy and Astrophysics Supplemental Series*, 107:243-264 (1994).
50. Loiseau, S. "Optical design for the Global Astrometric Interferometer for Astrophysics," *SPIE Proc.*, 2478:269-279 (1995).
51. Mahajan, V. N. "Zernike annular polynomials and optical aberations of systems with annular pupils," *Applied Optics*, 33,suppl:8125-8127 (1994).
52. Malbet, F., et al. "Space adaptive optics coronagraphy," *SPIE Proc.*, 2478:230-238 (1995).
53. Malbet, F., et al. "Active optics and coronagraphy with the Hubble Space Telescope," *SPIE Proc.*, 2201:1135-1144 (1994).

54. Mather, J. C. and B. D. Seery. "The Next Generation Space Telescope," *SPIE Proc.*, 2807:98-105 (1999).
55. ming Dai, Guang. *Theoretical Studies and Computer Simulations of Post-Detection Atmospheric Turbulence Compensation*. PhD dissertation, Lund University, Sweden, 1995.
56. Musinski, D., et al. "Metrology for the flight instruments for the Hubble Space Telescope first servicing mission at NASA/Goddard Space Flight Center," *SPIE Proc.*, 2478:167-174 (1995).
57. Needels, L., et al. "Limits on adaptive optics systems for lightweight space telescopes," *SPIE Proc.*, 1945:176-184 (1993).
58. Noll, R. J. "Zernike polynomials and atmospheric turbulence," *JOSA*, 66:207-211 (1976).
59. Oppenheim, A. V. and J. S. Lim. "The importance of phase in signals," *IEEE Proc.*, 69:529-541 (1981).
60. Paxman, R. G. and S. L. Crippen. "Aberration correction for phased-array telescopes using phase diversity," *SPIE Proc.*, 1351 (1990).
61. Paxman, R. G. and J. R. Fienup. "Optical misalignment sensing using phase diversity," *JOSA-A*, 5:914-923 (1988).
62. Paxman, R. G., et al. "Joint estimation of object and aberrations by using phase diversity," *JOSA-A*, 9:1072-1085 (1992).
63. Paxman, R. G., et al. "Evaluation of phase diversity techniques for solar-image restoration," *Astrophys. J.*, 466:1087-1099 (1996).
64. Press, W., et al. *Numerical Recipes in FORTRAN, The Art of Scientific Computing*. Cambridge, UK: Cambridge U. Press, 2nd Ed., 1992.
65. Primot, J., et al. "Deconvolution from wave-front sensing: a new technique for compensating turbulence-degraded images," *J. Opt. Soc. Am.*, 7:1589-1608 (1990).
66. Rao, S. S. *Engineering Optimization, Theory and Practice*, 3d Ed.. Wiley-Interscience: New York, 1996.
67. Roddier, F. "Curvature sensing and compensation: a new concept in adaptive optics," *Appl. Opt.*, 27:1223-1225 (1988).
68. Roggemann, M. C., et al. "Widening the effective field of view of adaptive optics telescopes by deconvolution from wave-front sensing: average and signal-to-noise ratio performance," *Applied Optics*, 34:1432-1444 (1995).
69. Roggemann, M. C., et al. "Comparison of Fourier phase spectrum estimation using deconvolution from wavefront sensing and bispectrum reconstruction," *Opt. Comm.*, 133:381-392 (1997).

Vita

Captain David Lee entered the U. S. Air Force in 1987, after graduating from North Dakota State University, with honor, in Electrical Engineering. He served with the 1000 Satellite Operations Group at Offutt AFB, NE as a satellite engineer for the Defense Meteorological Satellite Program, until coming to AFIT in 1992. At AFIT he completed a Master's degree in the Graduate Space Operations program of the AFIT Graduate School of Engineering, specializing in signal processing and image reconstruction topics. Continuing at AFIT, he is pursuing a doctorate in the Engineering Physics program, specializing in imaging physics. The Lee's will next be moving to Albuquerque, where David will be working at the Phillips Laboratory Starfire Optical Range facility, on Kirtland AFB, NM.

Permanent address: 1014 Middy Dr
WPAFB, OH 45433 .

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE April 1997	3. REPORT TYPE AND DATES COVERED PhD Dissertation		
4. TITLE AND SUBTITLE EVALUATION AND APPLICATION OF SPACE TELESCOPE ABERRATION SENSING USING PHASE DIVERSITY			5. FUNDING NUMBERS	
6. AUTHOR(S) David J. Lee, Capt, USAF				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology, WPAFB OH 45433-6583			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/DS/ENP/97-05	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Dr. Rich Carreras Phillips Laboratory, 3550 Aberdeen Ave SE KAFB, NM 87117			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; Distribution Unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Due to mechanical aspects of fabrication, launch, and operational environment, space telescope optics can suffer from unforeseen aberrations, detracting from their intended diffraction-limited performance goals. This dissertation gives the results of simulation and theoretical studies designed to explore how wavefront aberration information for such "nearly diffraction-limited" telescopes can be estimated via the Gonsalves (least-squares) phase diversity technique. In traditional phase diversity, aberrations are estimated from two images—one conventional, focused image and one image that has been defocused by a known amount. These studies are original and significant in that this effort represents the first time the effects of photon noise have been addressed for this relatively new aberration sensing technique. The effects of image-domain noise such as photon noise are of interest since the aberration estimates are derived from image-domain data. The analysis of this dissertation consists of simulation, theoretical investigation, and experimental application. Simulation studies incorporated models of the proposed Next Generation Space Telescope, as well a more generic, Zernike-aberrated model. Theoretical investigation involved deriving and numerically analyzing the Cramer-Rao lower bound for the phase diversity problem. The practical application dealt with diagnosing aberrations for an adaptive optical, ground-based telescope system.				
14. SUBJECT TERMS phase diversity, aberration sensing, space telescopes, adaptive optics, Next Generation Space Telescope			15. NUMBER OF PAGES 167	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	